nature

Arms control, SDI and Wall Street

Last week's abortive meeting in Moscow about the proposed agreement on intermediate missiles may not be too depressing. Both sides may learn from their contretemps and may yet sign the agreement.

MR George Schultz, one of the most amiable of men, has evidently had a rotten trip to Moscow. First, there was fog, which required that he should arrive by train. Second, knowing that his boss, the President of the United States, was even then (but not before time) going back on his nearly seven-year promise not to increase taxes in the United States, Mr Schultz had nevertheless to insist in his talks with Mr Mikhail Gorbachev that the same president would never abandon his conviction that the Strategic Defense Initiative (SDI) is the next best thing to an olive branch carried by a dove. Mr Schultz is probably not much better placed than the rest of us to know whether President Reagan will be able to hold to this line until the end of his presidency just over a year from now, but last week's events must have sown doubt in his mind.

On the face of things, several confusing things have happened all at once. On one reading of events, in the year since the abortive Reykjavik summit, the Soviet Union has lulled the United States into believing it would sign an agreement on the abolition of intermediate-range missiles, but, when the agreement was all but ready for signing, has imposed a post-condition — the requirement that there should also be an understanding about SDI. It is surprising that so many people are so surprised; this, after all, is what happened at Reykjavik. Indeed, Mr Gorbachev has been as consistent in his public declarations that SDI is an abomination as has been Mr Ronald Reagan in his insistence that only SDI can make the world safe from nuclear war. Is it not time that the two people came to an understanding of why they appear to differ so radically?

Time to sign

To hope to do that, they have first to sign the treaty on intermediate missiles, INF for short, whose negotiation is complete. If they fail to bring the INF treaty to a conclusion, they will find they have lost the goodwill and even interest of those states likely to be most affected by the treaty, those of western and eastern Europe. The Soviet Union has the most to lose in this respect, if only because it has gained most from its apparent willingness to give up its provocative deployment of its force of SS-20 missiles aimed at western Europe for the past decade and from its enthusiasm for the idea of a treaty. To fail to sign would also enhance the free-floating ambitions of the governments of Britain and France, each of which would sense a vacuum into which to flow. The general appreciation that Mr Reagan needs a treaty if he is to make it into the history books and that Mr Gorbachev will eventually come to see that a refusal to sign will be a political setback is one of the reasons why people have not been more downcast by Mr Schultz's trip to Moscow.

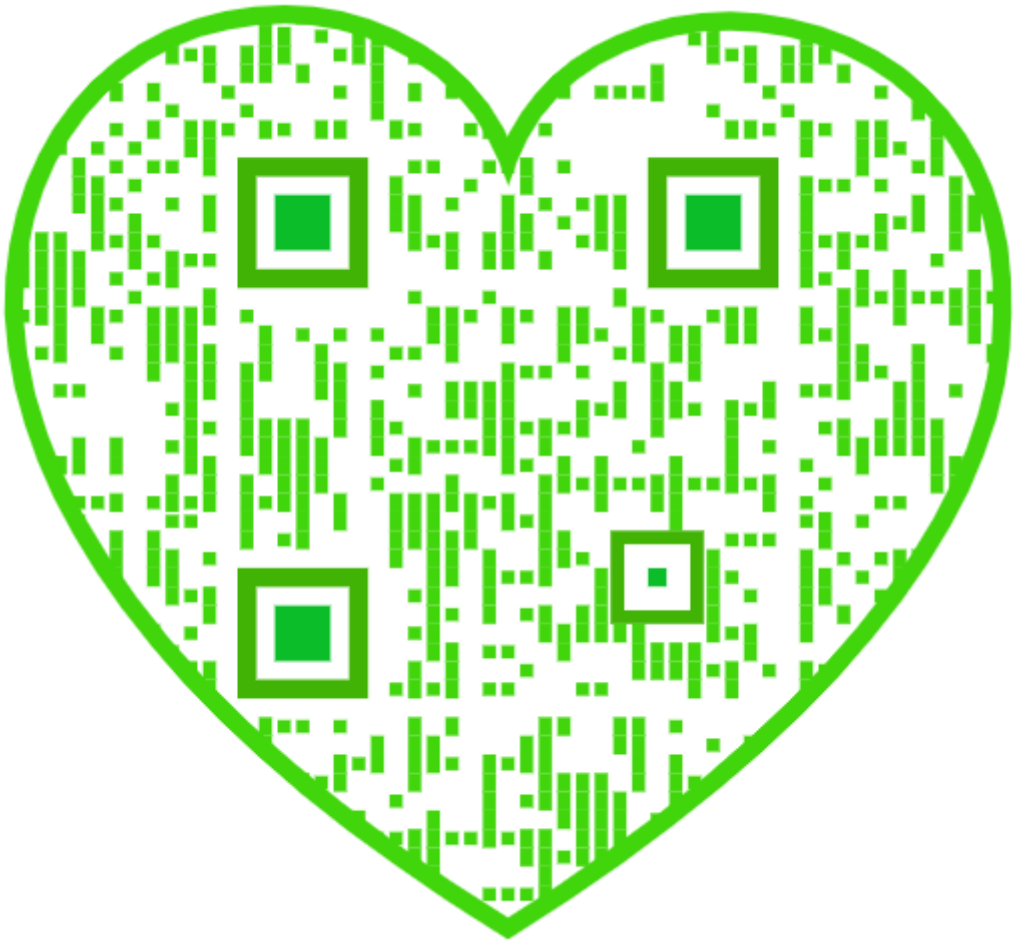
The questions remains of what should be said and done about SDI. The present impasse, illustrated by what Messers Schultz and Gorbachev said to each other in Moscow, is a consequence of calculated misunderstanding. The United States, or rather its president and his close advisers, says that SDI is a means of protecting whole nations from attack by hostile nuclear missiles, that a power sensing itself immune from attack would be able to let its offensive nuclear missiles rust in their launchers and that everybody would be able thenceforth to sleep more easily in their beds. Given the premise, the conclusion is in a probabilistic sense correct. Mr Gorbachev, on the other hand, says that there is no reason why a nation convinced it is immune from attack by other people's missiles should not also keep its own in good trim, planning to use them when a suitable occasion arises to knock the other fellow out. That is also an accurate calculation. Both people are logically on sure ground, and may continue shouting "yaa-boo" at each other for the rest of time unless something is done.

Bluff

The logical escape from each side's determination to see only one side of the other's case is for one to call the other's bluff. Mr Gorbachev is best placed to make the argument simply by asserting that moves towards the deployment of SDI should in logic, and for the sake of equal balance, be compensated for by reductions of the numbers of strategic missiles actually deployed. It is a simple point. If SDI is in any way effective, does that not imply that its users should be prepared to make unilateral reductions of their strategic missile forces? The same argument has often been used the other way about, as when supporters of SDI have said that adversaries would find it ruinously expensive to equip themselves with a countervailing force of offensive missiles. Does it not follow from that that an effective SDI should allow a nation to go without strategic missiles of any kind? Or to make do with, say, half a dozen?

At this point, the argument becomes academic. If it were indeed possible to develop an effective SDI (and even the enthusiasts acknowledge that the point has not yet been proved), those responsible would want some strategic credit for their cleverness. But the argument in its simple form shows quite clearly that the diplomatic way to neutralize SDI is in the context of negotiations about asymmetrical reductions of strategic arms. Why has a sensible person such as Mr Gorbachev overlooked that simple conclusion?

The simple answer is what people in the United States are often saying, that Mr Gorbachev's objective is not so much security (which matters nevertheless) as the more sensible use within the Soviet Union of scarce economic resources. If this is indeed the case, an INF agreement would bring very little relief. Virtually autonomous as they are, the Soviet military would not save much money for the civilian government if only one range of weapons were put on the shelf. That may explain why Mr Gorbachev so often talks in grander terms, of agreements to abolish nuclear weapons permanently; then at least it might be possible to think of bringing all those people (and their reliable production techniques) back into the civil economy. But that dream is a long way off. Maybe the most disappointing lesson to be learned from last week's meeting is that Mr Gorbachev has realized how slow will be the economic benefits of arms control, which may mean that he pays less attention to the possibilities. The best hope is that the events of the past few days on the world's stock exchanges will have pushed President Reagan in the opposite direction, towards a recognition that SDI (whatever its promise) is a poor use of economic resources. Is that too much to hope for?



Met. Office rides out the storm

- Once in a lifetime gales cause friction
- Met. Office set up internal inquiry

London

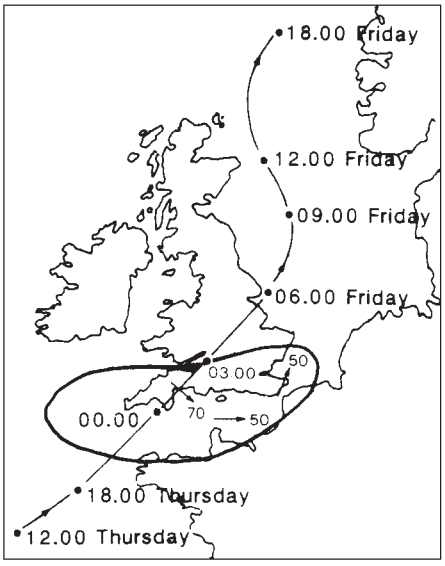
BRITAIN's Secretary of State for Defence, Mr George Younger, has ordered an independent inquiry into the failure of the Meteorological Office to forecast the storms that caused extensive damage on the night of Thursday–Friday, 15–16 October. The director-general of the Meteorological Office, Dr John Houghton, has already set up an internal inquiry but the Ministry of Defence has appointed two external experts, Sir Peter Swinnerton-Dyer, chairman of the University Grants Committee, and Professor Robert Pearse, head of the meteorology department at the University of Reading, to examine the findings of the internal inquiry and report to the ministry.

There has been much made of the 'Met. Office' failure to predict the storms — and claims that the European Centre for Medium Range Weather Forecasting (ECMWF) had fared much better — but is this criticism fully justified? Mr Martin Morris, assistant director of central forecasting services at the Met. Office has outlined some of the problems faced by the forecasters during the week of 15 October. A shortage of *in-situ* data certainly hampered the forecasters, but they were also faced with conflicting predictions from their models.

A wide range of measurements are used to forecast the weather. These include a network of surface and upper-air observations of temperature, pressure, wind, humidity and cloudiness. Reports from ships and aircraft are also available, as are satellite data. A large number of observations are still received for 00:00 and 12:00 GMT. For this reason, forecasts start from these times. However, significant numbers of observations are received for 06:00 and 18:00 GMT and satellites, aircraft and drifting buoys report at any time.

What were the medium- and short-range weather forecasts saying about the weather conditions on 15–16 October? The diagram shows the actual track of the depression. At 12:00 GMT on Thursday, the centre of the low was at around 46.5° N, 8° W in the Bay of Biscay with strong winds on the continental (south-east) side and a pressure of about 970 mbar. The Met. Office 12:00 GMT Tuesday forecast for 00:00 GMT Friday predicted the track of the depression well, but did not have the timing right (see table and figure). On Wednesday, the ECMWF model did not forecast the depression at all. The Met. Office fine-mesh model at 00:00 GMT on Thursday took the depression up the Channel, over East Anglia to the North Sea by 12:00 GMT (Friday) with little wind over inland areas but gales over the Channel and coastal areas. In contrast, the coarse-mesh model, run at a later time with more up-to-date data, produced an almost perfect forecast, with the low just south of Plymouth at 00:00 GMT on Friday at 966 mbar but the winds were not strong enough. With all this conflicting information, the forecaster compromised between the results of the fine-mesh and coarse-mesh models and had difficulties forecasting the track and speed of motion of the depression and the wind strength. Forecasts on national and local radio, and television forecasts in the south, warned of strong winds — but they were not emphasized on national television.

The low moved north much more slowly than expected and so its intensification and increasing wind strength were not predicted. As it moved north, however, data from the observation networks alerted the Met. Office and warnings of severe weather were issued to the emergency services. Met. Office scientists were slightly aggrieved by claims that French and Dutch meteorologists had forecast the storms better as ECMWF and Met. Office data are used by the French and Dutch and their warnings of stormy weather over the Channel were apparently based on the Tuesday Met. Office and ECMWF medium-range forecasts for Thursday and Friday. The Met. Office itself issued warnings on Tuesday to its weather centres for the commercial services.



Actual track of depression, and area of 40-knot mean wind speed at 06:00 GMT on 16 October. Some maximum mean wind speeds (knots) are given — a 70 knot wind is Force 12.

The information given by both the ECMWF and the fine- and coarse-mesh models was not consistent and so caused problems. The quality of the data used to specify the initial conditions for the models is of importance and the lack of a weather ship in the Bay of Biscay certainly did not help as little information could be gained about the strength of the winds. The weather ship *Romeo*, which used to occupy this station, was the victim of cuts last year. In the 1970s, there were eight weather ships in the west Atlantic; there are now only two. But as the models gave different forecasts it is not easy to say whether the missing weather ship data would have made much difference. It would have improved the initial data set but not necessarily the forecast. A careful look at the data received by these models is needed to assess why such conflicting forecasts were obtained.

Philippa Lloyd

Forecasts from Tuesday onwards for 15-16 October

Forecast	From	For	Description of low	Actual
ECMWF	12:00 GMT Tuesday	12:00 GMT Thursday	~47°N.23°W 968 mbar	~46.5°N 8°W 971 mbar
Met. Office (coarse-mesh)			~47°N 11°W 970 mbar	
ECMWF	12:00 GMT Tuesday	12:00 GMT Friday	Low over Ireland	East of Scotland 955 mbar
Met. Office (coarse-mesh)	12:00 GMT Tuesday	00:00 GMT Friday	Took low to ~56°N 00°W at 960 mbar	Just south of Cornwall ~956 mbar
ECMWF	12:00 GMT Wednesday	12:00 GMT Thursday	Low not forecast	
Met. Office (fine-mesh)	00:00 GMT Thursday	12:00 GMT Friday	Took low up Channel, over East Anglia into North Sea (957mbar)	
(coarse-mesh)	00:00 GMT Thursday	00:00 GMT Friday	Low south of Plymouth at 966 mbar	

Star wars X-ray lasers were oversold, says Woodruff

San Francisco

PHYSICISTS Edward Teller and Lowell Wood misled United States policymakers with "overly optimistic and technically incorrect statements" about X-ray laser research for the Strategic Defense Initiative (SDI), according to Roy Woodruff, former director of the X-ray laser project at Lawrence Livermore Laboratory, Woodruff left his directorship in 1985 because, he says, he knew he was not able to deliver what Teller was promising.

The allegations, made in a confidential grievance letter sent by Woodruff to University of California (UC) president David Gardner last April, were leaked to the press just as Woodruff's own dispute with the Lawrence Livermore Laboratory appears to be over (see below). Woodruff, a keen supporter of SDI, regards the release of the confidential letter, by the Southern California Federation of Scientists (SCFS), an anti-SDI group with which he has no connection, as being "totally inappropriate".

In the letter, Woodruff claims that laboratory director Roger E. Batzel knew of the inaccuracies presented to government officials by Teller, co-founder and former director of the laboratory, and Wood, a laboratory division head, but "refused to transmit correcting technical information or allow me to do so."

Because the information is classified, Woodruff's letter does not reveal the specific criticisms he made. But some Livermore scientists say Teller and Wood vastly underestimated the time necessary to perfect an X-ray laser system, and met with President Reagan and members of Congress on several occasions to promote the research and lobby for increased funding for SDI.

SCFS decided to release the letter now, according to Robert M. Nelson, co-chairman of SCFS, because budget decisions on SDI funding are being made in Congress. Nelson recalled President Reagan's speech last October, after the collapse of the Reykjavik summit, in which he said scientists were convinced that a space-based anti-weapon system could be deployed "several years down the road." If that was the President's thinking, said Nelson, "I would have been unfaithful to my responsibility as a citizen, not to make this information known."

Nelson said SCFS hopes for an outside investigation by an "appropriate agency," and it may get its wish. Woodruff was called on to give a classified briefing to interested members of Congress, on 23 October.

An aide to representative George Brown Jr, who requested the briefing, said a congressional investigation may ensue, to evaluate whether Congress has been operating under erroneous assumptions about SDI. **Marcia Barinaga**

Weapons researchers' right to dissent upheld

San Francisco

IN an action that backs the right of US nuclear weapons laboratory employees to voice dissenting views, the University of California (UC) has settled a personnel grievance filed by Lawrence Livermore Laboratory scientist Roy Woodruff.

Woodruff, who resigned in October 1985 as director of the laboratory's X-ray laser programme, claims he has been treated as "an un-person in a windowless office" and has not been given assignments that make use of his skills.

The university, contracted by the government to manage the Lawrence Livermore and Los Alamos National Weapons Laboratories, refused to hear Woodruff's first grievance, filed on 3 April, in which he criticized Lawrence Livermore Laboratory director Roger Batzel for refusing to correct "overly optimistic" representations of X-ray laser research made to President Reagan by physicist Edward Teller (see above).

A second grievance, filed by Woodruff on 30 April, requested simply that the laboratory reinstate him in a position of responsibility commensurate with his scientific skills. A committee of 3 members of the laboratory's technical and management staff heard Woodruff's case, and UC President Gardner accepted their recommendation that Woodruff's status be improved.

In a letter to Woodruff dated 20 October, Gardner said he was instructing the laboratory to find a suitable position for Woodruff within its defence programmes.

UC has been criticized for not taking a more active role in managing the weapons laboratories (see *Nature*, 329, 275; 1987), and Gardner's move has some bearing on the debate. Woodruff claims his treatment over the past year has been a form of reprisal for his disagreement with Batzel.

He believes the freedom to dissent is essential to the laboratory's role in advising the administration, and pointed to a recent paper by Livermore scientist Christopher Cunningham, that cast doubt on the viability of kinetic kill weapons (see *Nature* 328, 657; 1987). Marcia Barinaga

AIDS drug gets green light

Washington

THE controversial drug ribavirin, developed by ICN Pharmaceuticals for the treatment of AIDS (acquired immune deficiency syndrome), has been put back on track by the US Food and Drug Administration (FDA), in a surprise reverse decision. The FDA has approved clinical trials in humans of ICN's ribavirin, tradenamed Virazole, despite the strong reservations voiced earlier by the agency that the drug was inefficacious and produced harmful side effects.

Ribavirin is a guanosine analogue that ICN says has shown promise in preventing the progression of lympadenopathy syndrome (LAS) to full-blown AIDS in those infected with the AIDS virus. The drug has been controversial since its development, and the FDA has twice rejected ICN's application to enter clinical trials this year. At the Third International Conference on AIDS in Washington, DC in June, FDA Commissioner Frank Young stood up at the end of ICN's presentation of early clinical data, and issued a statement that the FDA had not approved ribavirin for use in humans, and that ICN had yet to give the agency evidence of the drug's effectiveness.

The FDA's reluctance to allow ICN to proceed with human trials of ribavirin has been criticized by gay rights groups and other AIDS activists as quibbling over details. In a move sympathetic to AIDS sufferers, the California state legislature passed a bill in September to side-step the FDA approval process and authorize tests of any experimental AIDS drug manufactured in the state on its own, including ICN's ribavirin.

ICN has drawn criticism from members of Congress and the FDA for trying to accelerate the approval process by releasing early data to the media. The company held a press conference last January to announce its data on the drug's effect in prolonging the progression of AIDS despite FDA suggestions that the conference be postponed until the results could be properly assessed.

ICN approached the FDA in its third application for clinical trials with a redesigned protocol and additional data on ribavirin's safety. In his announcement of approval of the drug for clinical trials, Young said the safety considerations were no longer of "sufficient magnitude" to prevent tests in humans from going forward. Although new FDA regulations permit drug manufacturers to sell drugs for life-threatening diseases during Phase II clinical trials, FDA has indicated that ribavirin may not be administered outside the clinical tests. **Carol Ezzell**

Merck's new drug free to WHO for river blindness programme

Washington

IF a newly-launched programme to eradicate onchocerciasis, or river blindness, a parasitic infestation affecting as many as 20 million people in Africa, south and central America and the Middle East, is successful, much of the credit will be due to Merck & Co., the New Jersey pharmaceutical company, which announced last week that it will provide its antiparasitic drug Mectizan free of charge to countries that need it. The burden of distributing the drug and supervising medical care falls to the World Health Organisation (WHO), but Merck was apparently the initiator of the idea and its driving force until WHO was persuaded of the potential effectiveness of the new treatment.

The parasitic worm *Onchocerca volvulus* is transmitted to man by the blackfly, which breeds in fast-flowing rivers and harbours immature worms, or microfilaria. In the human body, adult worms grow over a period of years, generating in turn more microfilaria which travel throughout the bloodstream. Infestation generally causes skin lesions and weight loss, but when the microfilaria travel to the eye, the resulting tissue damage causes blindness. About 340,000 people suffer from this extreme form of onchocerciasis, and in some Ghanaian villages 15 per cent of the population is permanently blind.

Two drugs, diethylcarbamazine and suramin, have been available for some time to kill microfilaria in the body, but both have severe side-effects and must be given often, making widespread use impossible in large and poor areas.

Control of onchocerciasis has thus centred on elimination of the blackfly, and a large-scale pesticide spraying programme, overseen by WHO, has operated since 1974 with some success. But without a treatment to kill the worms in the body, eradication seemed a forlorn hope.

Mectizan was developed by Merck in 1975 as a variant of ivermectin, a drug widely used against animal parasites. It kills microfilaria, and damages the reproductive system of the adult worms, putting a stop to infestation. Its great advantage is that it is effective in small doses given only once or twice a year, and has few side-effects. There may be adverse effects on pregnant and lactating women and small children, but Dr Mohamed Aziz, who supervised the clinical development of Mectizan, says the doses necessary to control onchocerciasis are about one hundred times lower than the level at which reproductive problems are seen in mice.

The idea of developing ivermectin for human use originated with Dr Bill Campbell, a veterinary doctor at Merck, Sharpe and Dohme who had lived in Ethiopia. After some 'arm-twisting' by Campbell and others, Merck agreed to begin work on clinical development. Aziz, who comes from Bangladesh, described the onchocerciasis programme as the "highlight of my life".

Eric Ottesen of the National Institutes of Health says that Merck deserves "an enormous amount of credit" for pushing the idea past some initial scepticism at WHO. The company will benefit by good public relations from the drug giveaway, but the project will certainly cost it money. Dr Roy Vagelos, Merck's chairman, said the cost would not amount to a significant fraction of Merck's running costs. Although Mectizan, because of its potency and safety, could in principle command a high price, Merck could not neccessarily make much profit from it because the only countries that need it have no money.

Until recently, US law would not have countenanced even the giveaway of Mectizan, as it has not been approved for human use by the Food and Drug Administration (FDA). But US law now permits export of unapproved drugs provided the recipient allows it, and in this case the countries concerned are unlikely to say no, especially as the French Directorate of Pharmacy and Drugs has approved Mectizan for human use.

Dr Halfdan Mahler, director-general of WHO, joined in the chorus of praise for Merck's "generous gesture", and hoped that it would help break down the "paranoia" that exists between WHO and the pharmaceutical industry. WHO has $4 million available to begin to set up the infrastructure for delivery, but will need more. Mahler hopes that the promise of a manifestly effective and relatively cheap programme such as this will persuade member countries that WHO is not the "bloated bureaucracy" that many think it. Mahler foresees visible reduction in the incidence of river blindness in a few years, and maintains that eradication of the disease by the end of the century is a reasonable goal. **David Lindley**

Groups of blind people being led by children. As many as 15 per cent of a village can be affected by river blindness.

Bhopal aftermath re-assessed

New Delhi

A REPORT just released by the Indian Council of Medical Research (ICMR) shows that fears of extensive damage to the health of survivors of the December 1984 tragedy may be misplaced.

Except for lung damage and associated disabilities, the methyl isocyanate (MIC) gas released from the Union Carbide pesticide factory had no major effect on survivors of the worst-ever industrial accident in India, which killed 2,000 people.

The studies were completed two years ago, but publication was withheld for fear of influencing the case now before the Bhopal court. The report is the first batch of ten research papers (out of the hundred or so received by ICMR) to be cleared by the government's legal experts. The remaining papers will be released later.

The ICMR launched some 30 investigations immediately after the accident. A major finding of these studies is lung damage and pronounced muscle weakness in many of the 10,700 people who needed hospital treatment. But lung function was severely impaired in only a few cases. One study noted that the number of "victims who had recognizable neurological syndromes was surprisingly low considering the magnitude of exposure to the toxic gas". Another study said that "any chromosomal aberrations found were minor". No mutagens were detected in urine, and no significant changes were found in human placentas. Suggestions that MIC could induce sterility in males were discounted by studies that showed no alterations in sperm count, mobility and morphology.

The ICMR studies did conclude, however, that long-term treatment might be necessary for people with lung damage who can no longer work. Prolonged psychotherapy is also proving necesssary for nearly a quarter of the survivors who are showing signs of anxiety and neurotic depression. K.S.Jayaraman

Nations meet to plan for world nuclear fusion effort

Munich

NATIONS from East and West took steps toward an unprecedented cooperation in nuclear fusion in an agreement signed in Vienna on 19 October.

Representatives from the Soviet Union, Europe, Japan and the United States agreed to meet in Garching, West Germany, for at least four months a year for three years, starting in April 1988, to coordinate planning for the International Thermonuclear Experimental Reactor (ITER). One purpose of such a reactor would be to determine if nuclear fusion is a technically feasible source of energy. Garching is the site of the West German Max Planck Institute for Plasma Physics.

The four partners will not decide until 1990 whether to build a common fusion reactor and, if so, where. Nevertheless West German fusion expert Siegfried von Krosigk of the Research Ministry said that the ITER agreement would give a "tremendous advantage" to all the partners, "regardless of whether a common reactor is built or not."

Willy Helin, a spokesman for the European Economic Community (EEC), said that the EEC agreed to participate provided that the Next European Torus (NET) is not limited in any way by ITER. NET is also supposed to demonstrate the technical feasibility of fusion power. Helin said that the EEC would welcome the possibility that an ITER reactor, if one is agreed upon, might be built in conjunction with NET. NET is not due to be constructed until the early 1990s and no site has as yet been selected.

Garching was selected as a meeting place over sites in Japan and the United States, and the process of elimination used may give an indication of how an international fusion reactor, should one be built, may some day find a home.

The Japanese site at Naka was eliminated because a new building would have had to be constructed and because the access to the international airport at Narita was poor. The US site at Lawrence Livermore Laboratories in California would have fulfilled the technical needs of the group — in particular scientific facilities already involved in fusion research — but Soviet insistence on a European site left only Garching.

The idea of a worldwide collaboration in fusion energy arose at the 1985 Geneva meeting of Mr Mikhail Gorbachev and Mr Ronald Reagan. The decision by Gorbachev and Reagan to plan the next steps in fusion research together has eventually led to the ITER agreement.

Steven Dickman

Relief at deal on Europe's research programme

Brussels

THE European Community finally achieved an agreement on its five-year Framework programme of research at last week's research council, but not without being reminded that the British government, largely responsible for the nine month's delay in settling the issue, retains the right to make further trouble at the end of the year.

The research council meeting on 28 September was sober, efficient and so quick that many of those present, expecting the adjournment which has been a tradition since the British presidency in the second half of 1986, were not there at the end.

What the council agreed was a budget of 5,396 million ECU (1 ECU = £0.69) for the period until the end of 1991, together with a carry-over sum of 1,084 million ECU. But 417 million ECU of the new money has been frozen, on British insistence, pending unanimous agreement on the European Communities' budget at the summit meeting fixed for Copenhagen in December.

Last week's council specifically agreed to spend 500 million ECU on the RACE programme of telecommunications technology in the five years beginning on 1 June as well as to commit 65 million ECU to medical and health research (including 14 million ECU for AIDS).

The Communities' information technology programme ESPRIT is also now in funds, although a detailed list of the projects for the second phase of the programme (called ESPRIT II) will not be approved until the next council meeting planned for 30 November.

Last week also saw a celebration in Brussels of ESPRIT's achievements so far at which Europe's West German research commissioner, Karl-Heinz Narjes, compared the treatment of the research budget unfavourably with various other spending items that "place an excessive strain on the budget"; while Jean Stern, who is the chairman of the French electronics company Bull, complained of the political climate that had compelled "a forceps birth" for ESPRIT II.

Bernard Conlon

Surrogacy agencies shut

Paris

MICHÈLE Barzach, the French minister for health and the family, has instructed the Keeper of the Seals to close within two weeks the three surrogacy agencies now operating in France. According to Barzach, the agencies — Alma Mater, Sainte-Sarah and Anias — contravene French civil law.

Under clause 351 of civil law, it is an offence for an individual to incite a woman to enter into an agreement to abandon her child. Dr Monique Vigy, writing in the French daily *Le Figaro*, points out that surrogacy cannot be described as a "scientific solution to infertility" — which would be legal. Neither the production of a sample of sperm by a man, nor its introduction into the neck of the uterus, could be defined as depending upon new medical techniques, Vigy argues. P.C.

International collaboration

London

THE Upjohn pharmaceutical company has embarked on a new collaborative venture with European academic institutions in an attempt to identify products with the potential for commercial exploitation. The first of a series of 'European Discovery Centres' was formally established last week between Upjohn, University College, London and the Centre Paul Broca de l'INSERM in Paris. Upjohn will support each centre for up to five years, with a staff of up to 20 people to carry out the research, which will be done in the host institution. Ultimately Upjohn hopes to spend around US $10 million annually to support work in ten discovery centres. Upjohn will have exclusive rights to inventions derived from the collaborations. S.L.H

Cornell site moved

Washington

AFTER stirring up a storm of controversy, Cornell University in Ithaca, New York, has agreed to move the proposed site for the new building that will house the Cornell supercomputing centre (see *Nature* 329, 280; 1987). The university's first choice for a site would have cut into the tree line above Cascadilla gorge, one of several gorges that slice through the campus. The new building plan calls for demolishing Grumman Hall and building the new structure on top of it, moving the building away from the gorge.

The old building was estimated to cost $29 million. The university expects the new design will cost $30 million. J.P.

Lawrence Berkeley: Correction

THE centres for sequencing the human genome for which the US Department of Energy has allowed in its budget (see *Nature* 329, 280; 1987) are at Los Alamos and Lawrence Berkeley National Laboratories. There is no centre at Lawrence Livermore National Laboratory. □

Greater financial accountability for UK higher education

London

As part of its overhaul of the structure and planning of higher education, the British government is proposing to introduce stringent financial controls over universities, polytechnics and colleges. Strict auditing and accounting procedures are to be implemented in an effort to demonstrate greater control over public funds and to avoid a repetition of the sort of financial mismanagement that occurred at University College, Cardiff, which earlier this year required £10 million of government cash to pay off debts (*Nature* **327**, 180; 1987).

Last week, the Department of Education and Science (DES) published a consultative document, *Accounting and Auditing in Higher Education*, emphasizing the need of institutions to demonstrate that "the resources made available to them are being applied economically, efficiently and effectively and for their intended purpose".

As part of its forthcoming education bill, the government intends to free polytechnics and colleges from the control of local authorities and to channel their funds through a Polytechnic and Colleges Funding Council (PCFC). The University Grants Committee will be replaced by a Universities Funding Council (UFC). An Education Assets Board (EAB) will be established, responsible for transferring assets to the polytechnics and colleges.

The government wants the Secretary of State for Education to specify to the funding councils and the EAB the information to be contained in accounts and how it should be presented. In turn, the funding councils will issue accounts directions to individual institutions.

Polytechnics and colleges will be required to follow present university practice and appoint external auditors. For all institutions the procedure for appointment and the appointment itself will, however, require the approval of the appropriate funding council. The funding councils will be required to "lay down a basic framework for the scope of the external audit". Such a framework the government foresees as including "whether institutions are complying with recommended practice; examination of accounting systems and internal controls and 'value-for-money' work". The 'value-for-money' monitoring is perplexing academics, who fear that the auditors will have the authority, against previously agreed principles, to encroach on academic judgments.

Within institutions, audit committees will be established, in broad compliance with a recommendation from a working group of the Committee of Vice-Chancellors and Principals. Such committees will comprise between three and six people, including senior lay staff, with the majority of the members of the governing body of the institution.

A temporary 'higher education internal audit unit' will be established for an initial period of 3–5 years to give support and guidance to the funding councils, and will be responsible to the accounting officers of the DES. The unit would have right of access to the books of the councils and the institutions they fund. At present the DES has no right of access to the books of individual universities.

The clear implication from the consultative document is that present arrangements for financial accountability are too loose. Although most of the proposals will not be made law, adherence will be required as a condition of funding. Academics say that while accountability for public funds is proper, such strict arrangements could serve to stifle local initiatives.

Simon Hadlington

Problems of Australian binary system

Sydney

THE joys of 'university' status seem to be proving too much for the binary structure of Australia's higher education system.

This week a second of the institutes of technology that provide the lower tier of higher education chose to be reshaped as a university. Others are anxious for change but will wait for the results of an inquiry into the binary system ordered by the new Minister for Education, Employment and Training.

The Australian binary system divides the 19 universities, which receive research funds from the Commonwealth government, from the 46 Colleges of Advanced Education, some bearing the title 'Institute of Technology', which do not. Only the universities are allowed to confer doctoral degrees and call their departmental heads by the title of professor. And the colleges are strictly bound by accreditation restrictions which prevent them altering their courses without government approval.

Less than a year ago, the West Australian Institute of Technology became the first college to change its name. Now the New South Wales state government intends to use its powers to reconstitute its own institute as the University of Technology, Sydney. The institute's administrators are blunt about the reason for the name change: it will bring money and prestige, particularly international prestige. They see it as unfair that while the quality of their research is often as good as that at the universities, their staff have to accept lower status.

Money will come largely from industry: the new name will provide opportunities for a company to endow a chair, for example. The Commonwealth government has no intention of letting the name change provide access to more public funds, however.

And the universities oppose any increase in their numbers for fear that it will mean that their research funds may one day have to be spread even more thinly.

Charles Morgan

Polish ministry is abolished

London

POLAND last week abolished its Ministry of Science and Higher Education. A new Ministry of National Education will now cover the entire range from kindergarten to postgraduate. What will become of the "science" sector is not entirely clear. Questioned on Warsaw Television on Monday, the incoming Minister of National Education, Henryk Bednarski, said that it would be necessary to reorganize and rethink the entire structure. One possibility would be to transfer the science sector of the old ministry to the State Committee for Science and Technical Progress, established in December 1984 by hiving off technology from the then Ministry of Science, Higher Education and Technology. The main purpose of this committee was to inject additional investment into research likely to bring short- or middle-term economic returns. But it is not entirely clear, however, what the status of this committee will be now. Its functions could well be subsumed, at least some of them, into the new Ministry of Industry.

Whatever happens to the science sector of the old ministry, however, Polish science will not be without at least one ministerial voice. The academic secretary of the Polish Academy of Science holds *ex officio* ministerial rank (in return for which, his first responsibility is to the prime minister, not to the members of the academy). Unlike the situation in the Soviet Union, Polish scientists cannot simultaneously hold university and academy posts, so those working in Academy institutes will not be affected by the shake-up. But the other two domains of Polish science — the universities and the "branch" (sectoral) research institutes, the situation, at present, seem somewhat unclear.

Vera Rich

Gloomy prospects for Britain's involvement in CERN research

London

PROSPECTS look grim for the high-energy physics community in Britain. According to physicists at a meeting in Geneva last week, a forthcoming report by Professor Anatole Abragam, of the Collège de France, will not meet British demands for substantial cuts in the budget of CERN, the European Organization for Nuclear Research. As a result, Britain may announce at the December meeting of the CERN council its intention of withdrawing from CERN in 1989.

In 1985 an inquiry commissioned by the British Science and Engineering Research Council (SERC) and the Advisory Board for the Research Councils (ABRC) and chaired by Sir John Kendrew proposed that Britain's membership of CERN should continue beyond 1989 only if it could do so "at a significantly lower cost". To achieve the necessary 25 per cent reduction by 1991 would require drastic cuts in the CERN programme, a new way of sharing the costs or the virtual abandonment of British experiments. None of these is reasonably possible.

SERC has already reduced its commitment to high-energy physics in response to pressure from other subject areas. An early victim of these cuts was the European Muon Collaboration at CERN which folded when the British contingent withdrew a year before the project was due to finish. According to Professor George Kalmus, director of the high-energy physics division of SERC's Rutherford Appleton Laboratory, Britain is missing out on some good physics. But the 'domestic' budget for high-energy physics has been shrinking faster than the 5 per cent per annum Kendrew recommends Britain now concentrates on projects at the proton–antiproton and Large Electron–Positron (LEP) colliders at CERN and HERA at DESY in West Germany.

Out of the total cost of high-energy physics in Britain of about £75 million this year, £58 million is the subscription to CERN. Spain's recently renewed membership and the prospective membership of Finland will do little to spread the cost of CERN. France and Switzerland have already rejected the use of a 'host's premium', as was proposed when the LEP project was initiated, to offset the purported benefits of having CERN sited in their territories. (They did, however, make additional material and financial contributions.)

Each member's subscription is based on its gross national product and costed in Swiss francs. The data used, however, are 5–8 years out of date. Although this has worked to Britain's disadvantage recently, the economic slump of the early 1980s will reduce Britain's future subscription. The weakening pound caused the 1987 subscription to jump from £38 million to £58 million. The Treasury is prepared to offset this additional cost for this year and next. Many at CERN think that SERC's budget should be permanently protected against currency fluctuations.

Nor are there any prospects of significant savings in CERN's budget, which is divided equally between materials and staffing. According to Herwig Schopper, director-general of CERN, there is "no possibility of saving money from the materials budget". Similarly, any savings from the cost of staff will not be apparent until well into the 1990s.

LEP, costing about SFr 1,200 million, has been built within a constant budget since 1981, and will cost SFr 80 million per annum to run when completed in 1989 — out of a total budget of SFr 774 million (in 1987 prices). Furthermore, maintenance has been held over because of the fixed budget. That CERN has continued to run is a testament, according to some physicists, to the high standard of engineering at CERN (often derided as gold plating).

Any cuts in staff numbers will initially cost more not less. CERN has already embarked on a programme to slim down staff by 10 per cent by the year 2000. The management was criticized in an interim report written by Professor Abragam, and Schopper accepts the criticisms. The interim report also called for the staff cuts to be accelerated, so that 300–500 jobs should be lost within two years.

Schopper suggests that future construction projects, such as the proposed Large Hadron Collider that would occupy the LEP tunnel, should be financed separately. The period of growth for high-energy physics, he says, is over. Indeed, between 1975 and 1980, the CERN budget was reduced by a quarter in real terms, and cuts "in other sciences are not resulting because of us". British scientists argue that rather than seeing the CERN budget in terms of SERC expenditure, it should be compared to the cost of an average-sized university, which similarly has an infrastructure to be maintained.

The consequences of Britain's withdrawal from CERN are unimaginable, says Schopper. Others suggest that the position of British scientists would be uncomfortable when cuts were being effected if Britain withdrew. One theoretician said the scientific loss would be greater than the financial loss. And before making a decision, Britain would do well to recall Mrs Thatcher's praise for CERN in 1982: "I think [it is] money well spent. The work will go on." **Roland Pease**

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Rutherford Appleton Laboratory.

A superconducting solenoid destined for the DELPHI experiment at the Large Electron–Positron collider at CERN. It started its tortuous journey from the Rutherford Appleton Laboratory near Oxford on Sunday 18 October. It is now waiting near Newbury, only 15 miles away, while highway, electrical and telecommunications engineers address the more important task of repairing Britain's utilities after the gales.

Scientific computing centre in Heidelberg

Heidelberg

THE University of Heidelberg has announced the creation of an Interdisciplinary Centre for Scientific Computing. The *Land* of Baden-Württemberg is expected to approve the project in November. If approved, the centre will bring in several computer scientists to join professors in the mathematics, chemistry, physics and biology departments whose research touches on computer-intensive problems.

The university is negotiating with IBM for the use of a vector-processing add-on for the existing IBM 3090 computer. Heidelberg mathematics professor Willi Jäger said that the centre will concentrate on the application of computers to problems such as gene sequencing, modelling the creation of galaxies, and the chemistry of combustion. There is already a concentration of researchers in Heidelberg working on stochastic mathematical modelling. Finding collaborators in computer science is a strongly felt need in the Heidelberg basic science departments in part because the university has no department of computer science.

The University of Heidelberg is trying to follow the examples set by the Courant Institute in New York and the University of Utah, albeit on a smaller scale. With the founding of an interdisciplinary centre in this field, Heidelberg is helping to create a trend in West German institutions toward breaking down barriers between the traditional faculties. Steven Dickman

Japan's scientific whaling not a commercial operation

Tokyo

JAPAN has been severely criticized for its plans to carry out 'scientific whaling' in the Antarctic during the coming austral summer. And, under threat of retaliation by the United States, the Fisheries Agency has drastically cut Japan's proposed whale catch for the next season. But what are Japan's plans for research, what effects is it likely to have on whale stocks and is it really a commercial operation in disguise?

Earlier this year, Japan proposed catching 825 minke wales and 50 sperm whales for 'research purposes'. But the proposal was immediately condemned by environmentalist groups as a device intended to skirt the International Whaling Commission's (IWC) moratorium on commercial whaling (see *Nature* **326**, 732; 1987). And at the IWC annual conference in Bournemouth (UK) in June, a US resolution was adopted calling for a scientific committee to review proposed research programmes.

Despite the IWC resolution, Japan pressed ahead with its plans, and in July the Agriculture, Forestry and Fisheries Committee of the lower house of the Diet unanimously decided to go ahead with the research. But the United States threatened to halve Japan's quota in the US 200-mile economic zone by invoking the Magnuson–Packwood Act, a domestic law designed to curtail the fishery catch of nations that ignore IWC recommendations; the US ambassador to Japan, Mike Mansfield, even suggested that Japan's proposal for scientific whaling would adversely affect deliberations on the omnibus trade bill now before Congress.

In the wake of the Toshiba affair, Japan had little choice but to retreat, and the Fisheries Agency has cut the proposed minke whale catch to 300 (no sperm whales are to be taken). This is a far cry from the 1960s, when Japan used to catch over 20,000 whales a year, valued at around ¥30,000 million ($200 million). But the irony is that Japan is not the country responsible for bringing some species of whale to the brink of extinction.

It was US and European whaling fleets in the 1930s that used to catch tens of thousands of blue whales in the Antarctic, severely depleting stocks. And it was US whalers at the turn of the century that brought the bowhead whale in the Bering Sea and Arctic to the brink of extinction. Current US estimates place the population of bowheads at only about 2,000. Nevertheless, the United States, through IWC actions, continues to support aboriginal subsistence whaling of 20–30 bowheads per year to try and preserve Eskimo culture.

In contrast, the IWC estimate the population of large adult minke whales in the Antarctic at 280,000 (430,000 if all sizes are included), while that of sperm whales was placed at 367,000 in 1979. And IWC estimates of the net annual addition of sperm whales (2,600 per year) far exceeds Japan's proposed take of 50 sperm whales. But the proposal has now been dropped in the face of US opposition.

The annual net addition of minke whales to the Antarctic population is uncertain. And, according to Kunio Arai, head of the Japan Whaling Association, one of the main puposes of Japan's research would be to determine natural mortality and pregnancy rates in the minke whale. Arai says that, under Japan's proposal, a new whale research institute will be established in the next few months with five scientists on the staff. Another 15–20 researchers from universities and research institutes throughout Japan will participate in the proposed programme for next season.

In 1947 a whale research institute was established in Tokyo with funds from the whaling industry, and the institute published around 400 reports on the biology and ecology of whales, many of them by leading cetologists, such as Masaharu Nishiwaki and Hideo Omura. But with the decline of commercial whaling the institute has been struggling to survive.

While it is debatable whether 20–25 researchers could handle 300 minke whales a year, it is absurd for the Western press and environmentalists to suggest that the Japanese proposal is a commercial operation in disguise. The costs of sending the 23,000-ton mother ship, the *Dai San Nisshin Maru,* with two catcher boats to the Antarctic for several months are enormous, and Arai says that with a catch of only 300 minke whales the operation will be well into the red; the balance will be covered by the Ministry of Agriculture, Forestry and Fisheries which has allocated ¥350 million ($2.4 million) for whale research in this fiscal year.

The US has expressed "appreciation" for Japan's revised plans in talks with Goroku Satake, director-general of the Fisheries Agency. But the decision will be left to the scientific committee of the IWC.

David Swinbanks

West Germans ground Hermes?

Munich

WEST German doubts about the proposed European space shuttle Hermes and the space station Columbus leave the three-legged European Space Agency (ESA) programme with just a single leg to stand on.

The West Germans, who are supposed to contribute 30 per cent to the costs for the Hermes shuttle and 38 per cent to the costs for Columbus, have made their reservations known privately in advance of the 9-10 November ESA ministerial meeting at the Hague. Total costs for the three projects — Hermes, Columbus and the booster rocket Ariane — are now estimated to surpass DM 30,000 million over the next ten years. The West German share will be more than 25 per cent.

Despite official protestations to the contrary, Ariane 5 is the only one of the three new ESA projects that the West Germans seem likely to support. The official government line still supports manned space exploration, despite what the Research Ministry calls "occasional criticism". Jürgen Rüttgers, a Christian Democratic member of the Bundestag Research and Technology Committee, said that Europeans will be shut out of space if they simply wait for US advances in rocket and shuttle technology.

But West Germany has serious reservations about both Columbus and Hermes. Columbus, though technically agreeable to West Germany, is hamstrung by doubts about its potential use by the US military. This is less of a problem for West Germany than for non-NATO members of ESA such as Switzerland, but West Germany is also put off by a perceived US wish to monopolize research results obtained on Columbus, even if the discoveries are made in a non-US laboratory within the space station.

The doubts about Hermes are even more serious, echoing the rumblings coming out of Great Britain (see *Nature* **329**, 660; 1987). Even Rüttgers said that the West Germans will reserve judgement on Hermes until all of the technical problems have been settled. These problems are said to be quite serious.

In a forceful speech at the West German Aerospace Research Establishment in Cologne on 21 October, French President Francois Mitterrand called upon West Germany for its unconditional support for Hermes. The speech illustrates the political dimensions of the struggle within ESA, and some West German officials see it as a potential blunder by Mitterrand.

The entire debate over doubling the costs of the three ESA programmes could hardly have come at a worse time for West German politicians, who are involved in an internal conflict over tax reform.

The three parties in the ruling coalition are due to meet on 3 November to discuss guidelines for the newly created West German space agency (see *Nature* **328**, 6; 1987). Then the full cabinet will sit on 4 November to determine the West German position for the ESA meeting the following week.

Steven Dickman

Field test for bright bacteria to go ahead soon

Washington

THE US Environmental Protection Agency (EPA) last week approved the first field test of a recombinant organism to fall under the Toxic Substances Control Act (TSCA). A unanimous vote by EPA's biotechnology advisory committee has cleared the way for Monsanto to field test a strain of fluorescent *Pseudomonas aureofaciens* containing *lacZY* marker genes that will allow the strain to be tracked in the environment. Monsanto plans to introduce genes for agriculturally beneficial proteins into the strain, and the marker genes will be used to assess the movement of the engineered bacteria.

TSCA is one of two regulatory statutes that EPA uses to control the release of recombinant organisms into the environment, under the Coordinated Framework for Biotechnology announced in June of 1986. It was originally developed for overseeing chemical substances, and requires a company or individual to notify EPA 90 days before making a new chemical not on EPA's master list of approved compounds. EPA normally conducts risk–benefit analysis to determine if the substance should be regulated, but if EPA fails to respond within the 90-day approval period, the company has tacit approval to make, use and sell the substance, causing concern that some notifications might slip through the regulatory net.

Most recombinant organisms are expected to be regulated under TSCA, but the first environmental release of a genetically altered bacteria — a *Pseudomonas* with a deletion for ice-nucleation, developed by Advanced Genetic Sciences — was permitted under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), considering ice to be a "pest".

Biotechnica International nearly won approval for its field test of a recombinant strain of *Rhizobium mililoti* containing genes for nitrogen fixation under TSCA this summer. But snags developed in Biotechnica's antibiotic tracking system, because the marker they intended to use was also found widely in natural bacteria. Biotechnica is now working on a new set of antibiotic markers, and expects to begin its field test in Pepin County, Wisconsin on alfalfa plants in the spring of new year.

The Monsanto field test is scheduled to begin in November, in Blackville, South Carolina.

Carol Ezzell

MIT needs millions

Washington

WINNING a Nobel prize clearly has its obligations. This week, the Massachusetts Institute of Technology (MIT) put its newest laureates, Susumu Tonegawa, who won the prize for medicine on October 12, and David Saxon, who won the prize for economics on October 21, on stage to launch an appeal for five hundred and fifty million dollars needed to ensure that MIT will continue as a world-beating institute.

The appeal, second in scale only to Stanford University's drive to raise a thousand million dollars, is the largest in MIT's history. A new science complex, to include a new biology building and remodelled facilities for the physics department — disciplines in which five of MIT's eight Nobel laureates work — is a major goal of the campaign. The new buildings and additional resources will allow a number of new research initiatives.

A nucleus fund of 210 million dollars has already been raised but finding the remainder in a five-year drive is not going to be easy with stiff competition expected from other universities seeking to renew outdated facilities. The funds should help MIT avoid too much reliance on short-term government research support which Professor John Deutch, MIT's chief academic officer, says "threatens to limit the range and character of the scholarly enquiry we are able to undertake".

Alun Anderson

Soviet-Chinese collaboration on the Silk Road

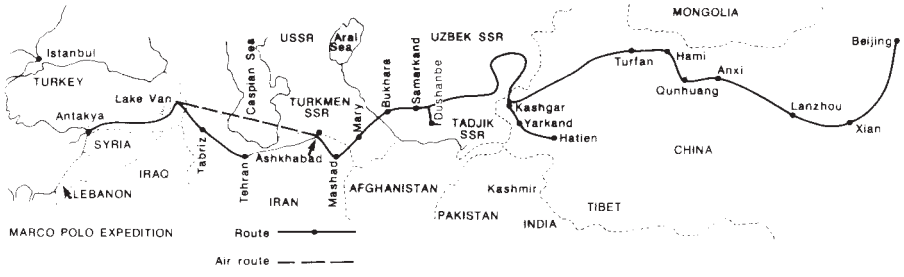
London

THE Marco Polo expedition, which set out in May to travel the entire length of the Silk Road from the Mediterranean to Beijing, proved a remarkable tribute to the current Soviet policy of *glasnost* and its Chinese equivalent. The five-man expedition, headed by historian Richard Fisher, hoped to be the first party to cover the entire route since the Portuguese monk de Goes in 1604. Whether they technically achieved this goal is uncertain — neither they nor their vehicle crossed the Iranian–Soviet frontier at the planned place.

In spite of the expedition having all necessary clearances from Tehran, the fronier guards refused to allow it to cross the 'closed' frontier, and eventually the expedition had to return to Turkey and enter the Soviet Union from there.

The team planned to investigate the ethnology and culture of the peoples living along the route, taking holograms *in situ* of cultural and archaeological relics and to study the siting of the traditional caravanserais, to investigate how far this was conditioned by cultural and how far by purely practical considerations. Dr Charles O'Neill of the Imperial Cancer Research Fund has suggested that silica fibres from wild grasses in flour along the Silk Road might be the reason for the high incidence of oesophageal cancer in those regions, and the team's programme called for the collection of flour samples from bakeries, grasses and epidemiological and *post mortem* data from the public health authorities. How far O'Neill's thesis could be tested today, when much Soviet wheat is pooled and consumed far from the site of origin is open to question. What is perhaps more significant is that, although some plant and flour specimens were collected, neither in the Soviet Union or in China were the public health authorities or the medical academies in touch with the expedition at all. Contacts arranged through the Royal Society were with the Academy of Sciences of the USSR and the Academia Scientia Sinica, which delegated liaison and cooperation to, respectively, their Institute of Geography and Institute of Geology. If multidisciplinary expeditions of this kind are to become more common in the formerly closed areas of Central Asia, the Soviet and Chinese hosts may find themselves expected to provide multidisciplinary assistance. This idea should appeal to Mr Gorbachev at least, since a main point of his restructuring programme for science is the elimination of inter-departmental boundaries. So far he has not attacked inter-Academy boundaries. The experience of the Marco Polo expedition suggests that this, too, may be necessary.

Vera Rich



The 'Silk Road' — last travelled in its entirety in 1604 by a Portuguese monk.

Harvesting human embryos

SIR—A social movement for the harvesting of embryos for medical, utilitarian and possibly even commercial reasons seems to be a real possibility. Recent discoveries[1,2] suggest that a number of neurological disablements may be alleviated, if not cured, by fetal brain grafts. Consider, then, the following scenario. An agent of a commercial corporation or a national agency approaches a young woman, preferably from a country of the European Economic Community, and offers her £15,000 if she will (1) become a surrogate mother, (2) have an abortion in England before the 23rd week of pregnancy and (3) donate the aborted embryo to a suitable foundation. The brain tissue can then be used to save or enhance the life of some eminent businesswoman or statesman or general, or even some with-it reproductive technologist.

Four objections to this frightening scenario must be noted and rebutted. First, that such actions would be illegal. In fact, the legal position on payment of money for surrogacy is at the moment uncertain. The mother has ownership of the aborted embryo but it is illegal for her to sell it. Gifting is not, I believe, covered. The crux of the legal objection seems to be the mechanism of payment and whether or not it can be brought to court.

Second, it may be said that such a woman would not be able to obtain an abortion. The fact is that in 1985, according to the Government Statistical Surveys Office of Population, Census and Survey, 171,873 abortions were carried out in England and Wales (this includes residents and non-residents). This does not suggest that there would be a great difficulty in obtaining such an abortion.

Third, it may be argued that there would be moral objections. However, a number of people, if they are consistent, should not be averse to the above scenario: for (a) the lives of valued members of the community are being saved or enhanced at the expense of those of unwanted embryos; (b) it extends and supports a woman's right to choose; (c) it conforms to current moral attitudes that value actual living persons' utilities more than potential persons' rights; (d) an embryo is being eliminated that is not capable of independent life.

Fourth and finally, it may be thought that with the number of aborted embryos available, there is no need to persuade or induce women to enact the above scenario. One does not know what the supply/demand match would be. However, suppose supply is greater than demand. Are the surplus embryos to be incinerated? Or are they to be sold or donated to other countries? Does the private sector have a role to play? Who owns the embryo? Should its donation be a matter of altruism, like blood? Or should it be a commercial enterprise? Or should the disposal of such embryos be banned?

It is because I think the above scenario not unlikely that I must reluctantly agree with the general tenor of Erwin Chargaff's article[3] and his profoundly pessimistic conclusions.

S. LAL

Department of Physiology,
Kensington Campus,
Campden Hill Road,
King's College London,
London W8, UK

1. Snyder, S. *Nature* **326**, 824–825 (1987).
2. *The Times*, 22 May (1987).
3. Chargaff, E. *Nature* **327**, 199–200 (1987).

Seeing eye to eye

SIR—In your tribute to Sir Peter Medawar (*Nature* **329**, 472: 1987), you quote his dismissal of the views of a nameless ophthalmologist on El Greco.

The ophthalmologist (Mr Patrick Trevor-Roper) demonstrated that when the paintings are viewed through a cylindrical lens of the right power and orientation, the oddity of the figures (and most of their aesthetic appeal) completely disappears. The speculation that the painter might have suffered from a corresponding astigmatism requires that the defect be unilateral. He could then view the subject with the abnormal eye, and the canvas with the normal one.

The hypothesis may be implausible, but it is not nonsensical.

A.C. DORNHORST

Department of Medicine,
Middlesex Hospital Medical School
and University College London,
Middlesex Hospital,
Mortimer Street,
London W1N 8AA, UK

Antimatter

SIR—In his review of antiproton annihilation (*Nature* **328**, 773; 1987), T. von Egidy is a little dismissive of possible practical applications of antimatter. Although the production of antimatter is always likely to be inefficient, there may be applications where utility justifies macroscopic production. Moreover, the problems of storing macroscopic quantities of antimatter, although difficult, may not be impossible. Egidy himself notes that this would require neutralization of the antiprotons with positrons and the use of strong magnetic fields and low temperatures, but these problems may have technological solutions.

Significant military exploitation of this technology is unlikely: the military have the ways and means to destroy us all anyway. Here, as elsewhere, the question of the military abuse of technology can have only a political solution.

On the other hand, there are applications to which antimatter may be uniquely suited. In particular it has a potential use as a rocket fuel for interstellar travel[1,2]. Interstellar travel with flight-times measured in decades is impossible with chemical rockets because chemical fuels are not energetic enough for the attainment of semirelativistic velocities. Matter–antimatter annihilation releases 6×10^9 times as much energy per unit mass as the reaction of hydrogen and oxygen, making interstellar flight possible with relatively small amounts of antimatter.

IAN CRAWFORD

Department of Physics and Astronomy,
University College London,
Gower Street, London WC1E 6BT, UK

1. Papailiou, D.D. *Frontiers in Propulsion Research*, JPL Tech. Memo 33-722, Jet Propulsion Laboratory, 1975.
2. Forward, R.L. *J. Brit. Interplanet. Soc.* **35**, 391 (1982).

No new biology

SIR—Because Schrödinger was an eminent theoretical physicist, his little book[1] was probably read by many physicists. They would have learnt just as much biology if they had read one of the then current popular biology books. But they would have felt no urge to do that. As H.J. Muller pointed out in a dismissive review[2] called "A physicist stands amazed at genetics", Schrödinger's book fostered the much-needed liaison between the physical and biological sciences. However, contrary to the impression Schneider[3] tries to convey, there was little or nothing in the book that was useful to biologists — especially biochemists. We already knew of the specificity associated with the linear array of amino acids in proteins, and those familiar with nucleic acids assumed a similar specificity in the arrangement of nucleotides. Calling such structures 'aperiodic crystals' did not clarify the issue.

The book raised a curious pseudo-problem about entropy. Schrödinger's cat, and his refrigerator before nuclear energy began to be used, depend ultimately on sunlight to pump out entropy. These comparable processes raise no philosophical difficulties: since the time of Lavoisier there has been general agreement among biologists that organisms are heat engines living in an abundant energy flux. Even the cleidoic egg is incompletely cleidoic, and life in the abyss depends on photosynthetically produced oxygen.

N. W. PIRIE

42 Leyton Road,
Harpenden, Herts, UK

1. Schrödinger, E. *What is Life?* (Cambridge University Press, 1946).
2. Muller, H.J. *Heredity* **37**, 90 (1946).
3. Schneider, E.D. *Nature* **328**, 300 (1987).

Will Britain stay within CERN?

The British government's decision about membership of the European high-energy physics consortium cannot be put off much longer. Britain should stay in, but spend nothing extra.

BEFORE the end of the year, probably before the end of November, the British government will be coming to a decision about the future of British membership of the European Nuclear Research Organization (really a high-energy physics laboratory) at Geneva. Even those who believe the British prime minister to have been needlessly beastly towards British academics will not envy her her task. It is a nasty decision to have to make.

Britain has been a member of CERN since the very beginning, more than thirty years ago. As with other international organizations, membership is regulated by a treaty that binds the partner governments together. It is technically possible to withdraw by giving a year's notice, but nobody has so far taken advantage of that clause. Like pulling out of UNESCO, that must be a solemn decision.

The possibility that Britain might withdraw was first openly canvassed two years ago by the Kendrew committee (see page 755), itself the product of prolonged glooming by the British research councils about the large proportion of the basic research budget being spent on high-energy physics. It has now largely been forgotten that the same issue was addressed in 1968, when Sir Michael Swann spelled out the difficulties of keeping science alive in Britain while spending money with the big boys on high-energy physics; he was unwise to have singled out biology as in need of special protection. With the decline in the value of sterling in the past year, the membership fee alone is roughly 10 per cent of total disposable research spending. Making full use of membership might cost as much again, but the research councils are deliberately cutting back on spending in this field and therefore do not let that happen.

Now it looks as if the compromise that might have made it possible for the British to continue as full members of CERN, the possibility of saving at least 15 per cent of the annual cost of running CERN, will not materialize. That is not surprising. CERN is embarking on the construction of its biggest machine yet, the electron accelerator called LEP. This can hardly be the time to ask it to slow down. Indeed, to do so is a kind of folly.

This is also, in any case, a time when other member governments are more enthusuastic about high-energy physics than Britain has been for several years. It can hardly make sense to ask that Italians, for example, who are just now Europe's cheerleaders of high-energy physics, should bottle up their enthuiasm to accommodate British cheeseparing. That, at least, is how it would seem to them.

In any case, Italians and others (such as the French) know that the British government cannot lightly withdraw from CERN. Doing so will only reinforce the impression now firmly established in Europe that the British are but second-rate Europeans. Why else has the British government spent two years (and two annual subscriptions) looking for a cheap way out?

Matters are now coming to a head because the end of the year is approaching. To be freed from its commitment in 1989, Britain would have to give notice at the December meeting of the CERN council. It may win a few weeks of reprieve if its partners agree that the issue would be more calmly discussed early in the New Year. Either way, there is not much time left, and little enthusiasm outside the high-energy physics community in Britain for continued membership. What, in the circumstances, should the British government decide?

The first thing to say is that the high-energy physics community has been shiftless in its advocacy of its objectives. (Shiftless, which means carelessly indifferent, should be replaced by neglectful, or even feckless.) We all know that it would be important if the top quark were found, because it would make life a little more symmetrical, and that the new machine at CERN may do the trick. But what would happen afterwards? High-energy physicists easily become indignant, but have little to say beneath the huff and puff.

The simple truth is that nobody quite knows what high-energy physics is for. People engaged with the most promising kinds of theories (called "string theories") say that they may be even more promising than had been believed, but that there is not yet a sufficient basis for telling which of their many predictions is likely to be the more correct (in the Popperian sense). Nobody should complain at that, for novel theories are notoriously difficult to interpret. Even so, the silence of those who normally proclaim the benefits of the next machine has been a little reminiscent of the dog that never barked.

That is why it seems most likely that the decision about continued British membership of CERN will be made on grounds entirely unconnected with the subject-matter of high-energy physics, with which politicians are not usually well-acquainted. It will instead be noted that the cost of British membership during 1987 will amount to £57 million, roughly 10 per cent of the disposable income of all the research councils put together, and that it would cost at least as much to make full use of the opportunity created by the CERN membership subscription. Especially because the research councils have not been persuaded in recent years to put up anything like that amount, their performance will be counted as a minus and membership of CERN will end.

That is what logic would dictate. The truth is that the British government will not relish being told off by its European partners for being a second-rate European, especially when it does not understand why first-rate Europeans (in the context of high-energy physics) are manifestly different. So, in the next few weeks, there will be an urge to find some different kind of compromise. This is what it could consist of.

Everybody concerned should first do a little arithmetic. The CERN subscription, now 10 per cent of total research council spending, is merely 5 per cent of the total fixed spending on academic research when the cost of university researchers' salaries is added in. Is that too big a burden? Many people in Britain who jib at 10 per cent would live happily with half as much. The snag, of course, is that there would be nothing left over with which to build the detector that might be transported to Geneva so as to recognize the new string particles that will eventually be described.

The simple diplomatic way of dealing with that awkwardness is to pretend it does not exist. If Britain said in December that it was proud to be a continuing member of the CERN consortium, and told only its own academics (when they asked) that there would be no extra money for research equipment, nobody on the European mainland would be especially downcast. Some might even be glad of the opportunity. That is why the British government's best course is to stay a member of CERN, but to let high-energy physics grants take their chances with others at the peer-review meetings.

John Maddox

Lymphocyte activation

How T cells communicate

James B. Breitmeyer

MOST circulating lymphoid cells are thymus-derived (T) lymphocytes, which are essential mediators of immunity to viruses and tumours as well as regulators of cellular and humoral immune responses. T cells recognize foreign antigens only in association with major histocompatibility products expressed on antigen-presenting accessory cells or target cells. Understanding the manner in which T cells interact with other cells is therefore essential to elucidating the general mechanisms of T-cell activation. Three papers on pages 840, 842 and 846 of this issue[1-3] present new structural information on two important cell structures that contribute to lymphoid cell–cell communication: CD2/T11 on the T cell, and LFA-3 (lymphocyte function-associated antigen 3) on the accessory/target cell.

Recent experiments using monoclonal antibodies and purified CD2/T11 or LFA-3 clearly establish that CD2:LFA-3 interactions participate in the process of thymocyte adherence to thymic epithelial cells, T-cell recognition of antigen-presenting cells, cytotoxic cell recognition of target cells, and erythrocyte rosette formation (reviewed in ref. 3). In this issue[3], Timothy Springer and colleagues show that LFA-3 exists in two alternative forms on lymphoblastoid cells, one sensitive to release from the cell surface by phosphatidylinositol-specific phospholipase C (PIPLC), the other insensitive to PIPLC but expressing a cytoplasmic domain sensitive to protease digestion when analysed in microsomal preparations. This implies that the large form (p29) is a transmembrane protein, whereas the smaller form (p25.5) is anchored to the plasma membrane by phosphatidylinositol linkage[4]. Fortuitously, molecular cloning studies confirm the two alternative forms of LFA-3, because complementary DNA clones corresponding to both forms of the molecule have been obtained. On page 840, Brian Seed describes[1] a cDNA molecule encoding the phosphatidylinositol (PI)-linked form of LFA-3. And Springer and colleagues now report[5] a second cDNA clone that encodes LFA-3 in the transmembrane form.

The deduced LFA-3 protein sequences for the PI-linked form and for the transmembrane form are identical from the amino terminus to residue 235, which corresponds to residue 207 of the mature protein[1,5]. The only area of divergence begins near the carboxy terminus of the membrane-spanning domain of the transmembrane (TM) form of LFA-3. Figure 1 shows the hydrophobic region common to both molecules and the regions of divergence. The calculated difference in relative molecular mass (M_r) between the TM and PI forms of LFA-3 on the basis of amino-acid differences alone is about 1,500. But the M_rs of deglycosylated mature LFA-3:TM and LFA-3:PI differ by 3,500 on SDS-PAGE analysis[3]. Although this electrophoretic estimation could be inaccurate, this result suggests that post-translational modification of LFA-3 could occur. Amino-acid sequencing of PI-linked variant surface proteins of *Trypanosoma brucei* show[4] that mature molecules lack hydrophobic carboxy-terminal amino acids predicted from cDNA sequences. A schematic representation of LFA-3 in which the TM form is expressed intact, and the PI-linked form undergoes post-translational modification with excision of the carboxy-terminal hydrophobic domain before addition of the PI–glycan moiety, is shown in Fig. 1. The resultant LFA–PI protein has a predicted M_r 3,500 less than the predicted TM form. This model suggests the two forms of LFA are identical in sequence from the amino terminus to the membrane-attachment site, consistent with the data of Springer and colleagues[3] showing identical amino-terminal amino-acid sequences for p29 and p25.5 molecules. Thus, a very minor variation in carboxy-terminal structure results in major divergence of post-translation processing for the two alternate forms of LFA-3, and may produce molecules with distinct functional properties.

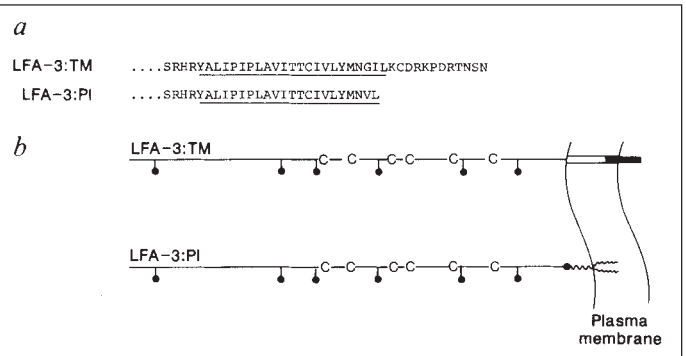


Fig. 1 The LFA-3 molecule exists in two alternative forms, differing only at the extreme carboxy-terminal end. *a*, Carboxy-terminal amino-acid sequences of the putative transmembrane form of LFA-3 (LFA-3:TM) (ref. 5) and the putative phospholipase sensitive form (LFA-3:PI) (ref. 1). Hydrophobic regions are underlined. *b*, Schematic representation of the two alternative forms of LFA-3. Cysteine residues (C) and potential N-linked carbohydrate attachment sites (↓) are shown for reference. The carboxy-terminal region where LFA-3:TM sequence varies from the LFA-3:PI sequence is shown in black. The carboxy-terminal hydrophobic domain encoded by the LFA-3:PI cDNA has been removed and replaced with a glycolipid membrane anchor (see text).

Using the ALIGN program, Seed finds that LFA-3 is very similar to its receptor CD2/T11, suggesting that LFA-3 is a member of the immunoglobulin supergene family[6]. Thus, the LFA-3:CD2/T11 system is strikingly analogous to recent data[7] on neural cell adhesion molecules (N-CAMs). In the case of N-CAM, alternate messenger RNA splicing results in both transmembrane and PI-linked forms of the molecule, although the alternate forms are more different from one another than are the two forms of LFA-3. More important, N-CAMs have recently been shown to participate in neural morphogenesis by affecting cell–cell interactions by homotypic adhesion. Thus, it seems that LFA-3 and CD2-T11 may have diverged from a common ancestral protein.

In addition to providing adhesive stabilization of cell–cell contacts, LFA-3 binding to CD2/T11 probably also contributes to T-cell stimulation. Pairwise binding of monoclonal antibodies to separate T11 epitopes provides potent polyclonal activation signals to T cells, clones, thymocytes and natural killer (NK) cells[8,9]. This result was originally interpreted as evidence of an antigen-independent pathway of activation, because NK cells and early thymocytes are activated through CD2 in the absence of T-cell antigen receptor (TCR) expression. At least three functional domains are recognized by anti-CD2 monoclonal antibodies[8,9]. On page 842 of this issue, Peterson and Seed present a new mutational analysis indicating that three highly limited regions of the CD2 molecule affect the determinants recognized by monoclonal antibodies. Mutations in each site abrogate the binding of a specific and non-overlapping group of anti-CD2 antibodies.

Analysis of extensive data on anti-CD2 antibodies reported at a recent workshop[9] produces some interesting conclusions about the functional domains defined by Peterson and Seed. Mutations of CD2 at residues Tyr 140 and Gly 141 abrogate binding of two antibodies that bind to activated but not to resting T cells, recognizing CD2 conformational epitopes[8,9]. Antibodies against these activation epitopes (anti-T11$_3$) are strongly mitogenic in combination with sheep erythrocytes, which express T11TS, the sheep homologue of LFA-3 (ref. 10). This suggests

that CD2 activation epitopes are essential in T-cell activation, but the physiological ligand for $T11_3$ remains undefined.

The antibodies whose binding is affected by CD2 mutations around residue Lys 48 are strongly mitogenic in combination with either anti-$T11_3$/1Mono2A6 or antibody 9.1 (S. Denning *et al.*; S. Yang *et al.*, in ref. 9). Although these mutations are generally associated with diminished erythrocyte binding in Peterson and Seed's studies[2], however, only some but not all the antibodies block rosetting[9]. Thus, residue 48 of CD2 defines a region of the molecule that is essential for pairwise antibody activation but may be peripherally involved with LFA-3 recognition. In contrast, the antibodies whose binding is affected by mutations around CD2 residue Gly 95 are generally not mitogenic with anti-$T11_3$ although they do uniformly block erythrocyte rosettes (D. Olive *et al.*, in ref. 9). Most interestingly, each of this group of antibodies is mitogenic with the 9.1 anti-CD2 antibody only in the presence of a poorly mitogenic anti-CD3 antibody, which binds to invariant chains of the TCR. This suggests that Gly 95 defines a region of CD2 essential for LFA-3 recognition, and which may also be involved in interactions between CD2 and the TCR complex.

Accumulating evidence suggests CD2: TCR interactions are essential in antigenic stimulation of $CD2^+CD3^+$ T cells. The combination of a poorly mitogenic anti-CD3 antibody and a single anti-CD2 antibody provides synergistic proliferative stimuli to T cells, suggesting that each moiety is providing a minimal signal but that the combination is fully activating[11]. Physical proximity of CD2 and TCR could contribute to this synergy, because crosslinking of TCR and CD2/T11 molecules by beads carrying both anti-T3 and non-mitogenic anti-$T11_1$ antibodies markedly potentiates T-cell activation (P. Anderson *et al.*, personal communication). Furthermore, stimulation of T cells by anti-$T11_2$ plus $T11_3$ causes potent activation of cellular kinases with phosphorylation of CD3 chains[12], suggesting that CD2 and TCR function are linked. In fact, in TCR-negative mutants of a human T-cell lymphoblastoid cell line, CD2 surface molecules cannot express the $T11_3$ activation epitope or generate intracellular calcium mobilization after anti-$T11_2$ plus $T11_3$ stimulation[13], suggesting that CD2 function depends on proximity of the TCR.

In all likelihood, then, LFA-3 adhesion to CD2/T11 and TCR recognition of antigen/MHC are synergistic rather than additive in stimulating T-cell activation during cell–cell contacts (see Fig. 2). LFA-3 binding to CD2 could cause activation (signal 1) if additional stimuli are provided to CD2 activation epitopes (such as $T11_3$). Complex formation between CD2 and the T-cell antigen receptor could cause synergistic activation signals (signal 2). Preliminary results[13] suggest that crosslinking LFA-3 on the accessory cell induces secretion of interleukin-1 (signal 3), although the data do not yet reveal whether transmembrane or lipid-anchored forms of LFA-3 mediate this signal. □

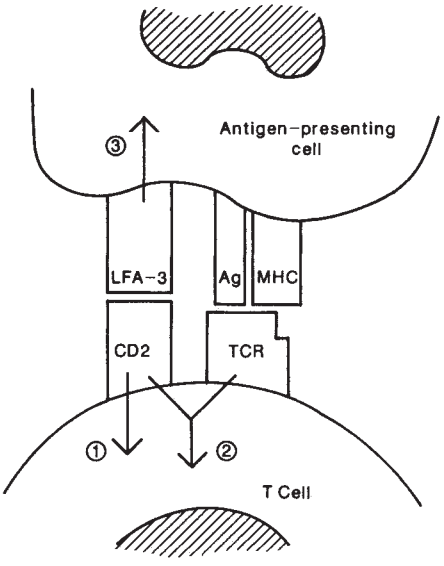


Fig. 2 Proposed intermolecular interactions during cell–cell contact between a T lymphocyte and an antigen-presenting or target cell. The TCR recognizes antigen in association with major histocompatibility complex products (Ag + MHC), whereas CD2 molecules interact with LFA-3. Potential activating signals resulting from these interactions include: (1) CD2 stimulation from LFA-3 binding in combination with a second signal delivered to the $T11_3$ epitope; (2) activation signals resulting from functional interactions of CD2 and the TCR; and (3) induction of lymphokine release in the accessory cell from LFA-3 perturbation. See text for details.

1. Seed, B. *Nature* **329**, 840–842 (1987).
2. Peterson, A. & Seed, B. *Nature* **329**, 842–846 (1987).
3. Dustin, M.L. *et al. Nature* **329**, 846–848 (1987).
4. Cross, G.A. *Cell* **48**, 179–181 (1987).
5. Wallner, B.P. *et al. J. exp. Med.* **166**, 923–932 (1987).
6. Williams, A.F. *Immun. Today* **8**, 298-303 (1987).
7. Cunningham, B.A. *et al. Science* **236**, 799–806 (1987).
8. Meuer, S.C. *et al. Cell* **36**, 897–906 (1984).
9. McMichael, A.J. (ed.) *Leukocyte Typing III* (Oxford University Press, 1987).
10. Hünig, T. *et al. Nature* **326**, 298–300 (1987).
11. Yang, S., Chevaib, S. & Dupont, B. *J. Immun.* **137**, 1097-1103 (1987).
12. Breitmeyer, J.B. *et al. J. Immun.* **139**, 2899–2905 (1987).
13. Le, P. *et al. Fdn Proc. Abstr.* **46**, 447 (1987).

James B. Breitmeyer is in the Division of Tumor Immunology at the Dana-Farber Cancer Institute, and at the Harvard Medical School, Boston, Massachusetts 02115, USA.

Natural selection

When learning guides evolution

John Maynard Smith

A PERENNIAL problem for evolutionary biologists is how natural selection can give rise to a complex structure that is of value to the organism when fully formed, but useless until it is complete. The orthodox answer, and I think often the correct one, is that it doesn't, and that complex structures have passed through a number of stages, each a little better than the last. For example, even random sequences of amino acids can have some, rather non-specific, catalytic activity. The vertebrate eye could not arise in a single step, but a single light-sensitive cell is better than nothing, a light-sensitive cell with a layer of pigment to one side is better still, and so on. That is to say, evolution is a hill-climbing process. In a recent paper, G.E. Hinton and S.J. Nowlan (*Complex Systems* **1**, 495; 1987) suggest how a combination of evolution and learning may make possible the evolution of a structure that is of value only when perfect.

An idea that has always had attractions is that adaptations acquired during an individual's lifetime, by learning or in other ways, are passed on to its offspring. For Lamarck, such inheritance of acquired characters was a main cause of evolution. For Darwin, natural selection was primary, but he also ascribed a role to "the effects of use and disuse". Weismann rejected the whole idea, essentially because he could not think of a mechanism whereby the structures acquired by an individual — for example, the blacksmith's muscles — could be translated into information in the gametes. With the development of modern genetics, and in particular of molecular biology, we have come to reject the idea that individual adaptation can alter the information in the gametes.

Even if we accept Weismann's view completely, however, it is still possible for individual learning to facilitate evolution. If individuals vary genetically in their capacity to learn, or to adapt developmentally, then those most able to adapt will leave most descendants, and the genes responsible will increase in frequency. In a fixed environment, when the best thing to learn remains constant, this can lead to the genetic determination of a character that, in earlier generations, had to be acquired afresh each generation. The idea goes back to J.M. Baldwin (*Am. Nat.* **30**, 441; 1896), C. Lloyd Morgan (*Science* **4**, 733; 1896) and C.H. Waddington (*Nature* **150**, 563; 1942). It has not always been well received by biologists, partly because they have suspected it of being lamarckist (a suspicion that Waddington was curiously reluctant to allay), and partly because it was not obvious that it would work.

What Hinton and Nowlan have done is

to answer these objections. They consider the following simple model, of whose biological unreality they are well aware. Imagine an organism with a neural net in which 20 connections must be correctly set ('on' and 'off') if its fitness is to be increased: if even one connection is wrong, it is no better than if all were wrong. This is the worst type of fitness surface, as there is no slope leading to the summit. There are 20 gene loci, each with three alleles, 0, 1 and ?. The first two specify on and off, and the third that the connection can be varied. During learning, an individual tries out a succession of settings for these variable connections. If, by chance, it tries out a correct set, and its genetically specified connections are also correct, it is rewarded, and does not alter its settings again. An individual makes 1,000 trials in its life. Its darwinian fitness is increased by a factor of $1+19n/1{,}000$, where n is the number of trials after it has learnt the correct settings. Thus, an individual that is born with the correct settings genetically fixed has a fitness 20 times that of one that never learns.

Hinton and Nowlan simulated a population of 1,000, with initial gene frequencies of 0.5 of ? and 0.25 of 0 and 1 at each locus. Because on average 10 loci were fixed, about one individual in 1,000 would have all the fixed settings correct. Also, with 10 variable settings, it would have a reasonable chance of learning the correct settings in 1,000 trials. Individuals were chosen as parents with probabilities proportional to their fitnesses. Reproduction was sexual. The response to selection was rather slow for the first 10 generations, because most individuals never learnt, but the response then accelerated, and after 20 generations the frequency of correct alleles was high, and learning was rapid.

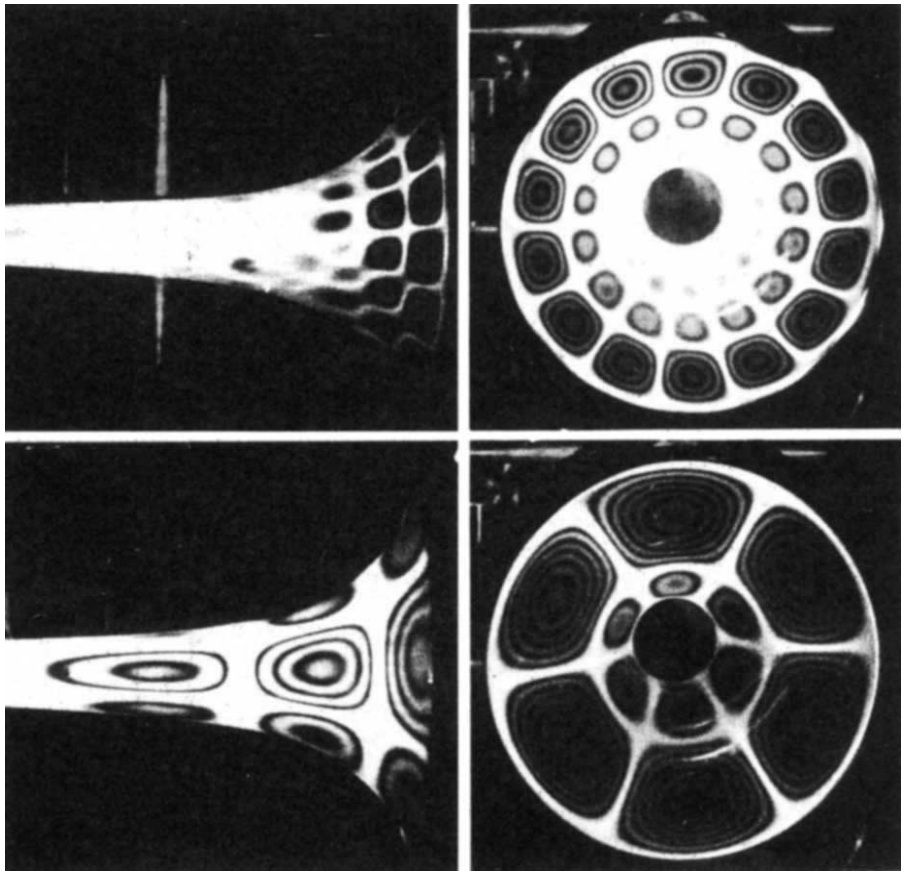
To what extent has learning accelerated evolution in this model? The absence of learning would be represented by the case in which only 0 and 1 alleles were present. There are then approximately 10^6 genotypes, of which one has a fitness of 20 and all the rest of 1. In a sexual population of 1,000, with initial allele frequencies of 0.5, a fit individual would arise about once in 1,000 generations (I ignore the fact that some wrong alleles would, in time, be fixed by genetic drift). Mating would disrupt the optimum genotype, however, and its offspring would have lost the adaptation. In effect, a sexual population would never evolve the correct settings.

What of an asexual population? An asexual population of 1,000 would be unlikely to include an optimal individual, and, unless one assumes an unreasonably high mutation rate, would never give rise to one. But a population of many millions would include optimal individuals that would breed true, and the correct settings would soon be established by selection.

Two comments are worth making. First, an asexual population, without learning, would try out more than 10^6 individuals before solving the problem, compared with 20,000 for the simulated sexual population with learning. Second, in the absence of learning, a large asexual population can evolve the adaptation, whereas a sexual one cannot (or does so excessively slowly). This illustrates the general fact that, when there are epistatic fitness interactions, sexual reproduction can actually slow down evolutionary progress, by breaking up co-adapted groups of genes as soon as they arise. (Epistasis arises in the present example because if, for example, A and B are correct alleles, and a and b incorrect ones, then A is fitter than a in the presence of B, but not in the presence of b.) Because sexual reproduction is almost universal, I wonder whether epistatic interactions can be as widespread as is sometimes thought.

Hinton and Nowlan show there are contexts in which learning (or developmental flexibility) speeds up evolution. It does so by altering the search space in which evolution operates, surrounding the optimum by a slope which natural selection can climb. To use their analogy, finding the optimal neural net in the absence of learning is like searching for a needle in a haystack. With learning, it is like searching for the needle when someone tells you when you are getting close. □

John Maynard Smith is Professor of Biology at the University of Sussex, Falmer, Brighton, Sussex BN1 9QG, UK.



THE shining brass of a trombone might look good, but does it have any effect on the sound produced? If the walls merely define the boundary of the vibrating column of air, one might just as well make trombones from plastic. On the other hand some musicians swear blind by their own favourite alloy. The figure shows the measured vibrations of a trombone bell played at 1,000 Hz (upper) and 630 Hz (lower). (Side-on views, left; end-on views right.) The holographs were recorded at the National Physical Laboratory, near London, as part of a study, by instrument designer Richard Smith, into the relationship between bell shape, material and sound. The contours reveal a distortion of the bell by a wavelength of laser light, about 0.5μm. The tests show that thinner bells vibrate more, and certain harmonics of some of the notes produced were up to 2 decibels stronger. Material effects were only apparent in thin-walled instruments. Blindfolded professional trombonists, however, were unable to hear the difference, and also were unable to distinguish the instruments in performance if they were adjusted for weight and balance. An electroformed pure-copper bell introduced in the blindfold tests went unnoticed, but mysteriously took on wonderful properties when the blindfolds were removed, copper being a 'superior' material. (Courtesy of Richard Smith, 110 The Vale, London N14 6AY, UK.) □

Superconducting ceramics

Now you see them, now you don't

J. Murray Gibson

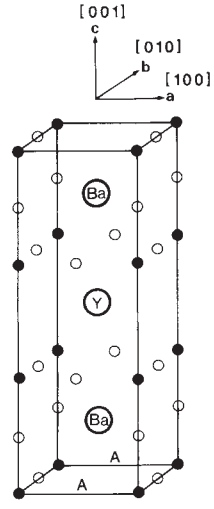
ARGUMENTS about the mechanism of high-temperature superconductivity often concern the arrangement of oxygen atoms in the crystals. More precisely, there is strong evidence that vacancies left by missing oxygen are ordered. However, controversy surrounds the ability of high-resolution transmission electron microscopy (HREM) to reveal the oxygen structure. On page 812 of this issue[1], N.P. Huxford *et al.* suggest that vibrations of the oxygen atoms will make them even less visible than hitherto realized.

Since the first HREM studies on the 90-K superconductor[2,3] $YBa_2Cu_3O_{7-\delta}$, people have speculated that the missing oxygen in the deficient perovskite structure is ordered within the unit cell and visible in the images. Copper–oxygen bonding in this material is believed to be the key to its superconducting properties, and so the ability to detect the vacancies is highly desirable. Neutron diffraction experiments[4,5] have clearly identified the location of the ordered oxygen vacancies within the unit cell of $YBa_2Cu_3O_7$. However the commonly used composition for a 90-K superconductor contains additional oxygen vacancies. Locating these and observing any variations in oxygen stoichiometry near boundaries and defects could be very important for understanding physical properties such as critical current. But it has been difficult, if not impossible, to identify clearly ordered oxygen vacancies in good single-crystal material using HREM.

The atomic arrangement of the superconductor is based on the cubic (unit-cell dimensions $a=b=c$) perovskite ABO_3. In this structure, metal atoms A and B sit at the corners and centre of the cubic unit cell, with oxygens at the centre of each edge. HREM images confirmed early diffraction work[6], showing that $YBa_2Cu_3O_{7-\delta}$ has a tripled unit-cell parameter in one direction, so that the structure is close to three unit cells of perovskite in a row $(CuBaO_3)(CuYO_3)(CuBaO_3)$. The stoichiometry of that unit would be $YBa_2Cu_3O_9$, so that there must be oxygen vacancies in the structure. A further clue to the arrangement of oxygen vacancies is given by the symmetry. A tripled-perovskite structure should be tetragonal $a=b\neq c$, retaining a square base. X-ray diffraction[6], however, shows that below room temperature the compound has orthorhombic symmetry (rectangular base, $a\neq b\neq c$). The difference in length between the a and b axes is very small and could be attributed to ordering in the oxygen vacancies. The definitive experiments that confirmed this ordering came from neutron powder diffraction[4,5].

Neutron diffraction is much better than X-ray or electron diffraction at identifying the position of oxygen atoms. The atomic-scattering amplitudes of X-rays and electrons are approximately proportional to atomic number, whereas neutron scattering amplitudes are roughly constant. Thus, scattering for 400-keV electrons from oxygen atoms is only 18 per cent of that from barium atoms at the angle appropriate for diffraction into the (100) Bragg reflection from the superconductor. In contrast, scattering of neutrons from oxygen is 75 per cent of that from barium atoms.

The structure of $YBa_2Cu_3O_{7-\delta}$ is shown in the figure. Neutron diffraction shows[4,5] that the oxygen vacancies are ordered in two ways: all the oxygen atoms at the same height (z) as the yttrium atoms are removed and half of the oxygen atoms around copper atoms between the barium planes are removed. It is the removal of the second set of oxygen atoms that imparts the orthorhombic symmetry, so that in the [010] direction one sees chains of oxygen vacancies in line, but in the [100] direction these oxygen vacancies overlap the copper atom positions. There is little doubt that these oxygen vacancies are of key importance to the superconducting properties of $YBa_2Cu_3O_{7-\delta}$. The oxygen vacancies around the yttrium atom impart a two-dimensional character to the extended electronic states and the additional ordered vacancies between the barium planes give a one-dimensional character to the electronic band structure and other properties. Although the exact role of the electronic, vibrational and magnetic properties in the superconductivity is still controversial, properties (for example the transition temperature) are dramatically influenced by the concentration and possible ordering of oxygen vacancies[7].



Tripled-perovskite (P_{mmm}) structure of the 90-K superconductor $YBa\,Cu_3\,O_7$. ○, oxygen atoms; ●, copper atoms. Note the absence of oxygen atoms in the plane of the yttrium atom and on the lowest plane (labelled A). The inset defines the unit-cell vectors **a**, ([100], 3.8198 Å); **b**, ([010], 3.8849 Å); and **c**, ([001], 11. 6762 Å).

Given that neutron diffraction is the best means of studying the long-range order of oxygen vacancies, why bother with HREM? One reason is that HREM allows spatial localization of structural data because the electron optics of a transmission electron microscope permit two successive Fourier transforms — from object to diffraction pattern to image. It has been postulated that variations in the oxygen stoichiometry and/or ordering are partly responsible for the relatively low critical currents observed for these superconductors. If oxygen vacancies can be seen in HREM images, then they can be spatially localized, for example, near grain boundaries.

Ourmazd *et al.*[8] presented the first experimental data combined with a careful set of image simulations to support claims that the HREM images are sensitive to the presence of oxygen vacancies. They identified an image feature that coincides with the row containing the vacancies between barium planes (A–A in the figure) as a 'fingerprint' of the oxygen ordering. According to their initial work, the best images of the vacancy fingerprint are obtained using thick samples near the Scherzer focus, at which the microscope aberrations are partially compensated for by a slight defect of focus, and the image is most directly related to the projected atomic potential. In thick specimens, however, multiple scattering of the electrons complicates this simple relationship. In thin specimens, where the simple relationship could exist, the image contrast from the oxygen vacancies is negligible: the change in average inner potential from the presence of oxygen vacancies is less than 2 per cent. In thicker specimens, the integrated scattering from the oxygen atoms grows in importance, but less rapidly than the multiple scattering from other strong diffracted beams. As a result, although images may become more sensitive to oxygen vacancies, they become very indirectly related to the real structure and extremely sensitive to many other experimental parameters that are hard to determine.

More recently, Ourmazd and Spence[9] specified the optimal conditions for strong contrast from the oxygen vacancies, indicating these difficulties. They concluded that it should be possible to distinguish the [100] and [010] projections in HREM images, but little else. Huxford

et al. in this issue[1] add another difficulty to the problem: the thermal vibrations of oxygen atoms as measured by neutron scattering. When these are incorporated in the diffraction and image simulations they reduce the intensity of the oxygen features to the point where most microscopists would consider them undetectable. But the thermal vibration amplitude measured by neutron diffraction could be confused for static disorder, which HREM imaging might be able to discern.

Unfortunately, it is hard to judge how reliable the fit of experiment and theory is. Although the algorithms for HREM image simulation take into account most important aspects of the imaging process, comparison of these results with experiment is left entirely to visual inspection. This controversy reveals the serious inadequacies of the qualitative approach, which does not allow scientists to make reasonable evaluations of experimental validity without being experts (if then). In contrast, we can readily assess the reliability of statistical fits to neutron-diffraction data[4,5] and have a simple yardstick of the accuracy. Because parameter space in instruments using two successive Fourier transforms exceeds that for diffraction, which involves only one, fits are even less likely to be unique. Statistical reliability must be addressed in ambiguous cases such as the vacancy fingerprint.

Nonetheless, electron microscopes have been useful in other ways for studying the new superconductors. The study of defects and multiple phases has been essential in isolating the perfect structure and certainly in discovering different superconductors. Convergent-beam electron diffraction is becoming better known as it can identify the full symmetry of the unit cell[10,3]. Dark-field imaging is tremendously useful in studying defects and twin boundaries[10]. Some defects initially detected and used to explain the high transition temperatures have since been found not to be intrinsic.

Lower-resolution diffraction contrast techniques could be better suited to studying oxygen ordering. For example, a Bragg reflection, such as (100), which is very sensitive to the oxygen vacancy concentration between the barium planes, can be used to maximize the signal-to-noise ratio in the image by excluding unwanted high-contrast beams. Furthermore, by tilting the crystal away from the symmetrical zone axis orientation in such imaging studies, the complexities of multiple scattering could be considerably reduced. An interesting example of this is a recent observation of superlattice diffraction spots probably arising from oxygen-vacancy ordering in $YBa_2Cu_3O_{7-\delta}$ with $0.3<\delta<0.4$ (ref. 11). Dark-field images with such superlattice reflections could be used to identify fluctuations in ordering across a sample at moderate resolution.

As often happens, we may have learned more about techniques than the problem in hand in studies of the superconductors. The fingerprint of oxygen-vacancy ordering in HREM images is at best very smudged, and were it not for other incontrovertible evidence, would probably be thrown out of court. □

1. Huxford, N.P. *et al. Nature* **329**, 812–813 (1987).
2. Syono, Y. *et al. Jap. J. appl. Phys.* **26**, L498–L501 (1987).
3. Beyers, R. *et al. Appl. Phys. Lett.* **50**, 1918–1920 (1987).
4. Beno, M.A. *et al. Appl. Phys. Lett.* **51**, 57 (1987).
5. Capponi, J.J. *et al. Europhys. Lett.* **3**, 1301–1307 (1987).
6. Cava, R.J. *et al. Phys. Rev. Lett.* **58**, 1676–1679 (1987).
7. Cava, R.J. *et al. Nature* **329**, 423–425 (1987).
8. Ourmazd, A. *et al. Nature* **327**, 308–310 (1987).
9. Ourmazd, A. & Spence, J.C.H. *Nature* **329**, 425–427 (1987).
10. Chen, C.H., Werder, D.J., Liou, S.H., Kwo, J.R. & Hong, M. *Phys. Rev.* B**35**, 8767 (1987).
11. Werder, D.J. *et al. Phys Rev* B (submitted).

J. Murray Gibson is Head of the Electronics and Photonics Materials Research Department, Physics Division, AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA.

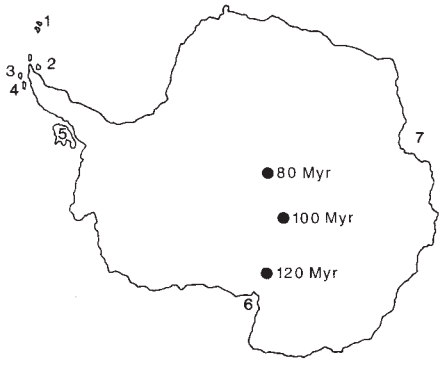
Palaeoclimatology

Explaining glacial periods

Eric J. Barron

THE most important aspects of the long-term evolution of the Earth's climate concern the observed alternations between cold, glacial periods and warm, non-glacial periods lasting millions of years. With the advent of continental-drift hypotheses, these great climatic shifts were attributed to changes in geography[1,2]. More recently, however, people have used the CO_2 greenhouse effect to explain the warm climates, and geographical mechanisms have been questioned. This change of emphasis arose largely because of results from climate models[3] and geochemical-cycle models[4]. In a provocative paper on page 803 of this issue[5], Crowley, Mengel and Short suggest that the shift in emphasis from geography to CO_2 was premature. Using a relatively simple climate model, these authors conclude that land area near either pole influences the seasonal cycle of temperature in polar regions sufficiently to control glaciation.

The greatest strength of the argument of Crowley *et al.* is its simplicity. Ocean and land surfaces have considerably different heat capacities. Consequently, large continental areas that experience seasonal variation in solar input have large seasonal amplitudes in temperature. Using an energy-balance climate model that successfully reproduces the modern amplitudes of seasonal temperature variation, the authors show that large land areas centred over a pole are characterized by high summer temperatures. These temperatures would be too high to permit glaciation. In contrast, if the pole is over a continent but near a coastline (within about 1,000 km), the ocean modulates the temperatures through the seasons. Crowley *et al.* suggest that cooler summer temperatures allow winter snow to remain on the land surface until the next winter season, hence glaciation is initiated. This argument is simple and direct, and too fundamental to be ignored. But three problems limit its power to explain the glaciated and non-glaciated climate states.



Positions of the South Pole 80, 100 and 120 million years ago. Also shown are the sites of Cretaceous deposits (1–7, from ref. 9) indicating that as the Pole moved across the continental margin, Antarctica was not glaciated, contrary to the argument of Crowley *et al.*[5]

First, almost all climate-model simulations to date use an uncoupled view of climate — climatic interactions of the atmosphere with the ocean or the cryosphere (snow, sea ice, glaciers and ice caps) are not included. Crowley *et al.* focus on summer temperature, but in the absence of ice-albedo feedback or a water cycle. Ice-albedo feedback (which occurs because temperatures are reduced as ice reflects solar radiation, leading to further accumulation of ice) certainly could influence the amplitude of the annual cycle of temperature. The water cycle, and seasonal variations in precipitation, together with ice-albedo feedback, must be a factor in the snow mass balance and in the probability of snow surviving a summer. The ocean heat flux and sea ice may also be involved. This criticism is not directed particularly at the paper by Crowley *et al.*, but reflects the status of modelling efforts. The specific question in the case of Crowley *et al.* concerns whether the summer temperature calculated in the absence of ice accumulation and ablation and in the absence of ice-albedo feedback is sufficient to explain the glaciation of the Earth. My view is that we need to couple the cryosphere and oceans to the atmosphere to answer these questions. Until that time, caution is warranted.

Second, the geological record offers many opportunities to compare pole position and glaciation. For example,

Crowell[6] presents evidence that glaciation moved across the ancient Gondwana supercontinent following the rotational South Pole deep into the continental interior. Glaciation ended when the South Pole moved offshore. To sustain the argument of Crowley *et al.*, it would have to be assumed that once an ice sheet was initiated it could shift and be maintained as the Pole position shifts into the continental interior. Figure 4 of Crowley *et al.* on page 806, however, shows that with the Pole in the centre of Gondwanaland, summer temperatures are too warm for glaciation. Furthermore, the Pole moved over Antarctica during the Jurassic and Cretaceous[6,7] (213–65 million years ago; see figure), a period noted for its warm, ice-free climate. This evidence also seems to contradict the model of Crowley *et al.* which asserts that glaciation should occur as the pole crosses the continental margin. In my view the record is always unkind to hypotheses that focus on one main cause for the climate record: exceptions can always be found, and these hint that several forcing factors are responsible.

Third, in all fairness, we must question the record of glaciations. The most recent comprehensive compilation suggests that glaciation could have lasted for much of geological time. The timing and extent of glaciation remains a problem. At a recent workshop on the long-term stability of the Earth system at Pennsylvania State University, several speakers suggested that the record of permanent ice is incomplete and that yet further periods could have been glacial. The correlations between cause and effect cannot be accepted too easily — more data are needed. The answers must involve models of ice-sheet growth coupled to models of the complete climate system, and should not depend on inference from the temperature of one season. Data from the geological record already offer many tests of the arguments and will probably yield exceptions to any hypothesis that relies fundamentally on a single causal factor. Crowley *et al.* make an interesting, provocative contribution. Even if we doubt that seasonality is the main factor, we must recognize the importance of the seasonal effect of continental distribution as one of the many factors in climate change. □

1. Crowell, J.C. & Frakes, L.C. *Am. J. Sci.* **268**, 193–224 (1970).
2. Donn, W. & Shaw, D. *Geol. Soc. Am. Bull.* **88**, 390 (1976).
3. Barron, E.J. & Washington, W.M. *J. geophys. Res.* **9**, 1267–1279 (1984).
4. Berner, R.A. *et al. Am. J. Sci.* **283**, 641–683 (1983).
5. Crowley, T.J. *et al. Nature* **329**, 803–807 (1987).
6. Crowell, J.C. *Trans. geol. Soc. S. Afr.* **86**, 237–262 (1983).
7. Ziegler, A.M. *et al.* in *Tidal Friction and the Earth's Rotation* (eds Bosche, P. & Sundermann, J.) 240–252 (Springer, New York, 1982).
8. Hambrey, M.J. & Harland, W.B. (eds) *Earth's Pre-Pleistocene Glacial Record* (Cambridge University Press, 1981).
9. Zeimeister, W.J. *Palaeogeogr. Palaeochim. Palaeoecol.* **59**, 197–206 (1987).

Eric J. Barron is at the Earth System Science Center, Pennsylvania State University, University Park, Pennsylvania 16802, USA.

Sociobiology

A darwinian theory of divorce

Jared M. Diamond

THE annual divorce rate in monogamous animals can be defined operationally as the percentage of pairs mated in one year, both of whose members survive to the next year and remate with new mates rather than with each other. In well-studied bird species, the divorce rate averages 21 per cent but varies greatly among species (Rowley, I. in *Mate Choice* ed. Bateson, P. 331–360; Cambridge Univ. Press, 1983). At one extreme, the rate is 50–73 per cent in migrant American songbirds which experience 50 per cent annual mortality, so that an individual returning to its previous territory in the spring cannot afford to wait and see whether its previous year's mate also survived to return. At the other extreme, long-lived albatrosses, which spend 1 or 2 years dating and swapping partners before settling on a mate, have rates of only 0–2 per cent. Two recent studies add to accumulating evidence that divorce in some cases is a darwinian adaptation to increase reproductive success.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Colour-banded oystercatcher returning to its eggs. (Courtesy of M. P. Harris).

The first study (Harris, M. P. *et al. Ibis* **129**, 45–57; 1987) involved a British population of oystercatchers whose marital statistics are reminiscent of Hollywood humans. The mating system of these birds is best described as serial monogamy punctuated by widowing or divorce: only one mate in a given year, but up to seven successive mates over the mean breeding life of ten years. No male divorces more than twice, but females divorce up to four times. The annual probability of divorce is highest (18 per cent) after the first year of breeding and decreases to 6 per cent for couples together for 3 years or more. As in humans, there is no clear relation between divorce rate and number of previous marriages, even when the first and fourth marriages of oystercatchers are compared.

It turns out that the best predictor of divorce is reproductive failure in the previous season. The divorce rate is highest for pairs that hatch or fledge no young, and decreases with the number of young produced. Thus, successful fledging of 2 or more young virtually guarantees continued marriage (only 1 divorce in 69 cases), but members of pairs that fledge no young frequently take new mates the next year (43 out of 251). The sole instance of divorce in the middle of the breeding season in this study promptly followed a nesting failure. Although reproductive failure in the first year of mating is disproportionately likely to lead to divorce, a failure in later years of a previously productive marriage has a greater chance of being forgiven. Hence the decision to divorce is based on reproductive history, with the most recent year receiving disproportionate weight.

The second study, of a colony of black petrels in New Zealand, involved smaller sample sizes but yields a similar interpretation (Imber, M. *Notornis* **34**, 19–39; 1987). No pair that succeeded in fledging a chick one season divorced in the next season, whereas one-third of the unsuccessful pairs divorced. As in the oystercatcher study, divorced males retain the pair's territory, and it is the female who moves out to another territory.

What is the adaptive significance of divorce? Reproductive failure can arise from the deficiencies of a mate, from the incompatibility of a pair or from poor cooperation. Hence, a member of an unsuccessful pair might increase its reproductive success by seeking a new mate. At face value, the oystercatcher data suggest that those hoped-for benefits of divorce are elusive, as males hatch or fledge no more young in the immediately post-divorce year than in the pre-divorce year, and a 61 per cent post-divorce improvement in fledging success for females fails to achieve statistical significance. But divorce, whether or not it is eventually beneficial, involves short-term risks, as newly formed pairs of birds tend to fledge fewer young than established pairs. At least for kittiwakes, patience following divorce is rewarded: breeding output declines for at least 2 years post-divorce, but in the long run it pays pairs that fail to hatch eggs to switch mates (Coulson, J. *J. anim. Ecol.* **35**, 269–279; 1966).

Thus divorce, like so many other behaviours, can be explained in darwinian terms. In fact, the criterion of proved reproductive failure that birds use for mate rejection is much more straightforward than the criteria for mate selection, which depends on characters

(such as a vigorous song) that may or may not correlate with reproductive success.

Do these studies of birds shed any light on divorce in humans? The parallels at first seem strikingly to suggest that our divorces are ultimately an adaptation to produce and successfully rear more babies. Human divorce is associated with various sorts of marital discord, which would have placed children at risk under the stringent hunter–gatherer conditions prevailing throughout most of human evolution. Middle-aged men tend to desert their grown children and menopausal wife for a younger woman, thereby enabling the scoundrel to father still more children. Men in some societies divorce wives who bear them no children or who bear them only daughters. But such men betray their biological ignorance: the sex of human offspring is determined by sperm rather than egg chromosomes, and male infertility contributes as much as does female infertility to barren marriages. Hence men who divorce for supposed darwinian reasons may only be running away from their own problems. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Substance K receptor

Return of the magnificent seven

Michael R. Hanley and Trevor Jackson

RECEPTOR cloning is rapidly becoming big business. Historically, it was necessary to implement a daunting programme of purification, structural analysis and rigorous functional assessment to pursue the structure of a particular receptor. But recent innovations in expression cloning may open the door to the systematic cloning of every receptor for which we have a name. On page 836 of this issue[1], Masu *et al.* report the molecular cloning and sequencing of the substance K receptor using only its functional activity in *Xenopus* (toad) oocytes. Its sequencing may present an example of a novel problem for the future: that primary structure (amino-acid sequence) becomes known in the absence of conventional characterization of receptor pharmacology and biochemistry. Moreover, the substance K receptor attracts special interest because it is the first receptor for a neural peptide to be cloned.

Substance K is a member of a family of structurally related peptides called the tachykinins, which share a spectrum of similar biological effects, including actions on sensory processing and inflammation. From several lines of evidence, it now seems likely that each of the three identified mammalian tachykinins may have a corresponding receptor.

The cloning strategy, described in detail by the authors, is conceptually straightforward. They constructed a bovine stomach complementary (c) DNA library in a lambda phage vector containing the SP6 promoter for *in vitro* transcription. The resulting mixture of cDNAs was used to generate a mixture of synthetic messenger (m) RNAs, which was then evaluated for expression of substance K receptor following injection into *Xenopus* oocytes, thus generating an mRNA library. The resulting mixture was taken through a series of fractionations until a single clone was obtained. The appealing aspects of this strategy are that only functionally defined, full-length clones are isolated, and that it should be applicable to any receptor or other gene product that can be assayed by oocyte expression.

It is immediately apparent from the predicted amino-acid sequence that the substance K receptor (SKR) is a single subunit that fits well into the 'seven-transmembrane segment' superfamily of M1- and M2- muscarinic receptors, β_1 and β_2 adrenergic receptors and the rhodopsin-related visual pigments. This designation is based on the prediction, common to each of these sequences, that the receptor is folded so as to cross the membrane seven times. The amino terminus contains two consensus sites for N-linked carbohydrate attachment (Asn residues 11 and 19), placing the first segment on the outside of the membrane. The result of these predictions is to generate a model structure as shown in Fig. 1, with four extracellular and four intracellular segments. The real amino terminus is unknown, but the sequence predicts that the amino terminus may be blocked by covalent fatty acid modification. In other membrane proteins, the Met–Gly sequence is an accurate predictor of amino-terminal conjugation of myristic acid to the Gly residue[2]. This is shown by drawing the first segment as a loop, tethered by the amino terminus. The intracellular segments do not contain any obvious consensus sequences for kinase action, but the authors call attention to a carboxy-terminal segment that is rich in Ser or Thr residues. There are no clearly defined tyrosine kinase sites.

The location of the ligand-binding site in SKR is unknown, but two suggestions have been made for the β_2 adrenergic receptor that may be useful to consider. Based on deletion analysis, the binding site has been suggested[3] to be analogous to the interaction of retinal with opsin, which occurs via transmembrane segment VII. Alternatively, crosslinking studies[1] direct attention to the third extracellular loop and transmembrane segment IV.

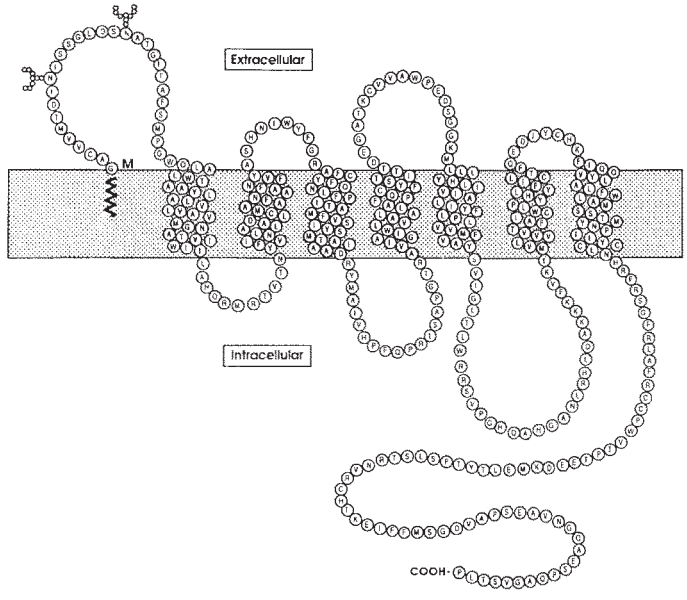


Fig. 1 Possible model for insertion of substance K receptor in the plasma membrane. Membrane-spanning segments are based on predicted hydrophobic transmembrane helices. The standard one-letter code for amino-acid residues is used. Also noted are the potential sites of N-linked glycosylation and myristylation at the predicted amino-terminal glycine.

It has recently been proposed[5] that peptides and their cognate receptors will show a strict complementarity at the nucleotide level. The cloning of the substance K receptor and the substance K precursor[6] from the same species permits an experimental test of the claimed predictive scheme. The predicted substance K binding site (H-termination–SNKGICFM) is not found in the substance K receptor. Thus, the complementarity hypothesis, reported to be successful in generating antibodies against leutenising-hormone releasing hormone and adrenocorticotropic hormone receptors[6] does not seem applicable to SKR.

The sequence similarity is greatest in the rhodopsin

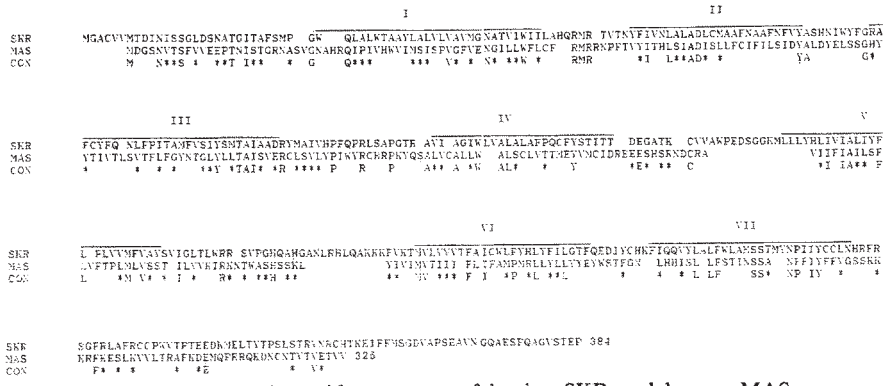


Fig. 2 Alignment of the amino-acid sequences of bovine SKR and human MAS oncogene product. Residues conserved in both proteins are shown below the sequences; the asterisk indicates favoured substitutions[7]. Positions of putative transmembrane segments I-VII are indicated.

family members and SKR. Using slightly different alignments from Masu *et al*[1] the degree of similarity, using Dayhoff rules for conservative substitutions[7] is SKR: M1, 38%, SKR: M2, 34%, SKR: β_2, 39% and SKR: rhodopsin, 46%. Masu *et al*[1] note that these results predict SKR would be coupled to a second-messenger transduction system by a GTP-binding protein. It is unclear what the best guess for the second messenger should be, but other evidence implicates the inositol lipid signalling pathways[8]. This poses an interesting dilemma, because there are three sequenced receptors linked to inositol lipid hydrolysis: platelet-derived growth factor (PDGF) receptor, M1-muscarinic, and now SKR, but they do not have any obvious intracellular site for the appropriate G protein. Indeed, the PDGF receptor belongs to a completely different class of receptor, having the extracellular domain, one transmembrane segment, and an intracellular domain with a tyrosine kinase activity. Thus, more sequences will be awaited with considerable interest in the expectation that functional consensus sequences may eventually be defined for interactions with the appropriate G-protein species.

As noted in an earlier News and Views article[8], the substance K receptor may be uniquely associated with mitogenic activity among the family of receptors to SK and related peptides. This suggests that we should be on the lookout for any similarities in sequence with identified oncogenes. As shown in Fig. 2, SKR has striking similarity to the human oncogene *mas*[9]. The *mas* oncogene was first isolated by Wigler and colleagues using total genomic DNA and a novel tumorigenicity assay, so it was predicted to encode a transforming receptor with an unknown ligand. When compared with the family of seven-transmembrane segment receptors, *mas* is most similar to SKR (44%). Consequently, SKR and *mas* could define a subfamily with growth-control activities. As a corollary of this, the *mas* oncogene product is predicted to couple to inositol phospholipid metabolism, a biochemical response repeatedly correlated with mitogenic activity[10]. □

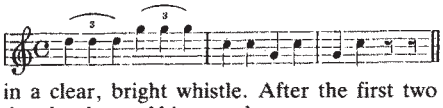
1. Masu, Y. *et al. Nature* **329**, 836–838 (1987).
2. Rein, A. *et al. Proc. Natn. Acad. Sci. U.S.A.* **83**, 7246–7250 (1986).
3. Dixon, R.A.F. *et al. Nature* **326**, 73–76 (1987).
4. Fraser, C. *et al. Proteins* **2**, 34–41 (1987).
5. Mulchahey, J.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9714–9718 (1986).
6. Nawa, H. *et al. Nature* **306**, 32–36 (1983).
7. Dayhoff, M.O. *et al. Atlas of Protein Sequence and Structure*, (ed. Dayhoff, M.O.) Vol. 5, Suppl. 3 (National Biomedical Research Foundation, Maryland, 1978).
8. Hanley, M.R. *Nature* **315**, 11–13 (1983).
9. Young, D. *et al. Cell* **15**, 711–719 (1986).
10. Berridge, M.J. *Biochem.J.* **220**, 315–360 (1981).

Michael R. Hanley and Trevor Jackson are in the MRC Molecular Neurobiology Unit, University of Cambridge Medical School, Hills Road, Cambridge CB2 2QH, UK.

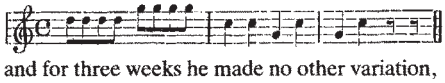
100 years ago

In *Nature* for August 11 (p. 343) there is an interesting article on music in nature; the writer evidently being inclined to deny that true musical notes, and especially several notes in succession . . . can be found in bird songs. However this may be in the Old World, we have in the New one example of a bird which not only sings, or rather whistles, pure and well-sustained musical notes, but has a succession of notes to form a simple melody. I refer to the scarlet tanager.

While we were at The Thousand Islands early in the summer of 1886, one of these brilliant fellows carried on a courtship among the trees close to our cottage, repeating incessantly during the first two days that we heard him sing the following strain,



in a clear, bright whistle. After the first two days he changed his song thus:-



and for three weeks he made no other variation, except that he occasionally repeated the last two notes a third time, thus filling out the bar. The notes were taken down by a trained musician, and if whistled gave the tanager's song exactly.

Coburg, Ontario. A.P. COLEMAN

From *Nature* **36**, 605; 27 October 1887.

Marine science

Ocean flux of Chernobyl fallout

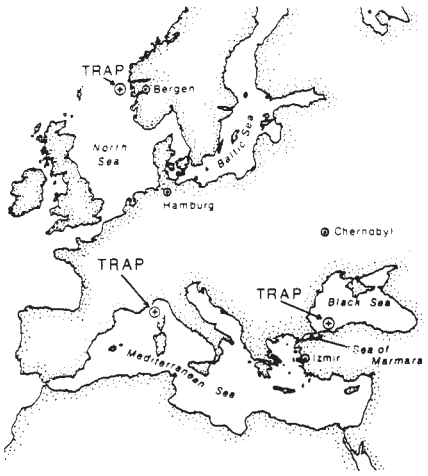
Michael P. Bacon

THE disaster last year at Chernobyl in the central Soviet Union is, somewhat surprisingly, of considerable interest to marine scientists. The accident effectively initiated a world-wide 'tracer' experiment on a scale that would never have been planned deliberately. K.O. Buesseler *et al.* and S. Kempe and H. Nies on pages 825 and 828 of this issue, and S.W. Fowler *et al.* in previous work (*Nature* **329**, 56–58; 1987), report the use of Chernobyl fallout to trace the downward transport of sediments in the Black, North and Mediterranean seas.

One of the most important developments in marine science over the past 10 years has been the growth in the use of sediment traps to record fluxes of material carried by particulate matter settling in the ocean. It is now possible to leave moored sediment traps unattended at any depth in the ocean for periods of 1 year or longer and to collect as many as 25 sequential samples of the settling particles at predetermined time intervals (Honjo, S. & Doherty, K.W. *Deep-Sea Res.*, in the press). This is the technique used to observe radionuclides from the Chernobyl fallout 1,000 metres below the sea surface in the Black Sea and 200 metres below it in the North and Mediterranean seas.

The observed flux of fallout at depth in the oceanic water column is a dramatic demonstration of the processes by which the oceans assimilate materials and become cleansed of contaminants. Marine plankton play leading roles in these processes. Primary production by phytoplankton creates the solid surfaces on to which reactive contaminants are adsorbed, and the feeding and digestive activities of zooplankton cause the aggregation of fine suspended debris into larger particles that sink rapidly (tens or hundreds of metres per day). Both the papers in this issue report that the largest fluxes of radionuclides occurred during plankton blooms.

It is of particular interest to note the speed with which the fallout nuclides first arrived at the collection depth. The traps



Map of Europe showing Chernobyl and the sites where Chernobyl-fallout radionuclides were collected with deep-sea sediment traps.

in both the North Sea and the Mediterranean Sea had, serendipitously, already been set for other purposes and had begun their sequential sampling at the time of the Chernobyl accident. In both cases it required less than 10 days for the pulse of fallout to reach 200 metres in depth, implying that the arriving particles had a sinking speed of at least 20 metres per day. That figure is within the range of values reported for the faeces pellets produced by zooplankton, and Fowler *et al.* argue that zooplankton faeces represent the principal pathway by which the Chernobyl fallout nuclides were transported downwards in the water column.

The idea that herbivorous zooplankton play an important role in accelerating the sinking of particulate matter in the oceans is not new. Evidence was first presented by C. Osterberg *et al.* (*Nature* **200**, 1276–1277; 1963) at a time when large releases of radionuclides into the atmosphere were a regular occurrence. Short-lived fission products from nuclear-weapons tests were detected in sea cucumbers living 2,800 metres deep off the coast of Oregon, requiring the existence of a mechanism that would deliver them to that depth in a matter of days or weeks.

The observations of the Chernobyl fallout are a direct demonstration of the operation of such a mechanism. It is now widely accepted that most of the vertical flux of particles in the ocean is caused by relatively rare large particles that are aggregates of one kind or another, and not by the more abundant fine-particle background. Much current research effort is aimed at understanding the origin and dynamics of large amorphous aggregates, often called 'marine snow', whose role in transporting materials downwards in the oceans is less well understood than that of the identifiable faecal pellets. □

Michael P. Bacon is at the Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

Therapy and the ideal of chemistry

"To analyse the patient — that is to separate his mental processes into their elementary constituents and to demonstrate these instinctual elements in him singly and in isolation". These are the words Freud used[1] to make an analogy between psychoanalysis and chemical analysis. It is immediately obvious that the analogy relates to technique — to separate and to isolate — and thus refers less to chemistry in general than to Lavoisier's definition of chemistry as an operational science and of the chemical 'facts' as those that have been 'purified' from all uncontrollable circumstances.

But the history of chemistry reveals that implementing Lavoisier's ideal so that chemistry could stop being an interpretative 'art' requiring the experience of a lifetime as described by the chemist Venel[2], and become a methodic practice where knowledge and action have a harmonious relationship, mainly occurred because nineteenth-century industry did produce the 'pure' reproducible products and the instruments such a practice needed. How could psychoanalysis make a similar break with the old interpretative art of traditional therapy to become a professional activity guided by a codified technique?

This problem is strongly connected with that of hypnosis. It is remarkable that Lavoisier himself was involved in the first discovery of the "power of imagination". He was part of the Royal Commission[3] which concluded that the universal 'fluid' by which the magnetist Mesmer explained the collective (hypnotic) crisis he provoked did not resist methodic investigation by separation and isolation. Indeed, the commission used 'placebo' methods to demonstrate that when imagination is purified away, the phenomenon disappears or becomes erratic. A century later (1885–1886), Freud learned from Charcot[4] that hypnosuggestion could give him a power like that of chemists: it could do and undo pathological symptoms just as the chemist composes and decomposes substances.

"Psychoanalysis proper began when I dispensed with the help of hypnosis", wrote Freud later[5]. Indeed, he had discovered that hypnosis did not give the power of control over his patients needed to verify the truth of his explanatory hypotheses, as they could lie or resist. Hypnosuggestion could not be purified from a complex of affective, unconscious factors which distorted the influence of the therapist. The solution Freud found for this problem is well known: he transformed this complex, the transference relationship, from an obstacle into a tool. It is the progressive elucidation of this relationship which must achieve the purification process, the production of actors able to verify without resistance the validity of scientific hypothesis.

Today, the problem of hypnosis has surfaced again. As Freud himself said[6], many analysts recognize that the ideal of conquering the resistance of patients, of dissolving transfer, is not operational. Instead of being purified away the affective, unconscious ties seem to be uncontrollably amplified by the psychoanalytical setting. In such a setting, the instalment of unrecognized hypnoid states could even be common. Contrary to Freud's hopes and to the example of chemistry, psychoanalysis has not succeeded in breaking completely with the old uncertain art of therapy.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Men at work — Freud (by Max Pollak) and Lavoisier (by P. Fouché)

Two centuries after Lavoisier, it is time for therapy to renounce the epistemological ideal of chemistry, of purification and isolation, to face the problem of what we call imagination or influence. This implies the meeting of many disciplines as it is clear that the problem is indissociably physiological, ethological, psychological and social. If freudian psychoanalysis has not succeeded in becoming a science, it is not because it is not rational enough, but because the relevant rationality is still to be invented.

Léon Chertok and Isabelle Stengers

1. Freud, S. *Lines of Advance in Psychoanalytic Therapy* Vol. XVII, 159 (Hogarth Press, London, 1919).
2. Venel, G.F. in the *D'Alembert and Diderot Encyclopedia*, t. III. (1753).
3. *Rapport des Commissaires Chargés par le Roi de l'Examen du Magnétisms Animal.* (Imprimerie Royale, Paris, 1784.)
4. Chertok, L. & Saussure, R. *The Therapeutic Revolution: From Mesmer to Freud* (Brunner/Mazel, New York, 1979).
5. Freud, S. *General Theory of the Neuroses* Vol. XVI (Hogarth Press, London, 1916).
6. Freud, S. *Analysis Terminable and Interminable* Vol. XXIII, 211 (Hogarth Press, London, 1937).

Léon Chertok is at the Institut de Psychiatrie La Rochefoucauld, 23, rue de la Rochefoucauld, 75009 Paris, France; Isabelle Stengers is at the University of Brussels, CP 231, Bd. du Triomphe, 1050 Brussels, Belgium.

Quantum effects in narrow channels

SIR—John Maddox's recent piece[1] "But how thin is a thin wire?", takes into account the size of the quantum effects on an electrical current conducted by a wire so thin that the current is affected by those quantum effects. A simple analysis of superfluids in narrow channels of a solid superleak, "Controversial drops" was also published[2] some months ago. In this article an experiment is suggested to analyse the nature of the outcoming drops from a solid superleak after the transit of the superfluid ^{4}He through the superleak.

Taking into account the quantum nature of the superfluid and that the size of a narrow channel of a solid superleak, in practice, can be some orders of magnitude smaller than a thin wire, I wonder if the two cases are similiar, if not identical. In both cases the same question arises: how narrow is the narrow channel of a solid superleak?

An interesting situation would be the case of a solid magnetic superleak[3] through which the magnetic phases of superfluid ^{3}He were in transit. In particular, A. Leggett (personal communication, 1977) said, ". . . if in these superleaks the adequate magnetic fields could be produced, quite a new kind of experiment in superfluid ^{3}He could be made, even its polarization". But, to my knowledge, no experiments with this kind of solid magnetic superleaks, with the magnetic phases of superfluid ^{3}He, have yet been made.

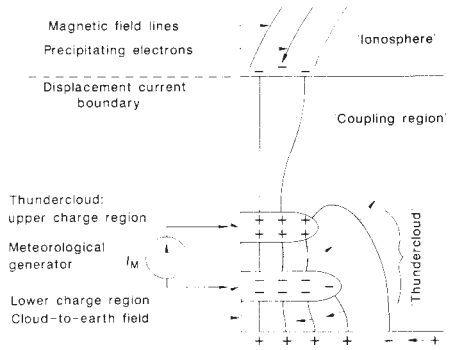
J.M. GOLDSCHVARTZ
Clavecimbellaan 273,
2287 VK Rijswijk (ZH),
The Netherlands

1. Maddox, J. *Nature* **328**, 197 (1987).
2. Goldschvartz, J.M. *Cryogenics* **26**, 692–693 (1986).
3. Goldschvartz, J.M. *Nature* **266**, 824–825 (1977).

Lightning triggering and synchronization

SIR—It has long been known that lightning causes magnetospheric 'whistlers', and the related precipitation of high-energy trapped electrons into the ionosphere has been demonstrated. Recently, Armstrong[1] has presented evidence suggesting that magnetospheric phenomena may influence the lightning which causes whistlers, at least in the sense that these phenomena may sometimes be synchronized with inter-hemispheric whistler echoes. Armstrong also reviews previous suggestions of how lightning may be triggered by external events. Both Armstrong[1] and Rycroft[2], in an accompanying introductory article, suggest rather complex triggering mechanisms whose difficulties are recognized by the authors.

A simpler mechanism for the triggered synchronization of ordinary lightning by



Capacitive coupling of charge from ionosphere to thundercloud. Negative electrons deposited on upper 'plate' cause field reconfiguration (indicated by transverse arrows) with more flux lines connecting from cloud-to-ionosphere, hence fewer lines from lower to upper charge but more to Earth, creating a larger cloud-to-Earth field.

precipitated electrons is suggested here, based on 'capacitive' coupling, in which displacement current couples charge variations from the ionosphere to the thundercloud. A simplified schematic of the fields and processes associated with a thundercloud is shown in the figure. A 'meteorological' charge separation process (I_M) maintains upper positive (Q_U) and lower negative (Q_L) charges. The top-to-bottom cloud field is charged directly, and the cloud-to-Earth field is charged via more circuitous paths, including the ionosphere and global circuit[3]. Breakdown is usually initiated from the lower negative charge, presumably due to exceeding a critical electrical field[4]. Precipitated electrons produce an ionospheric charge separation which deposits a negative charge, ΔQ, on the 'upper plate' of the cloud-to-ionosphere capacitance. (This displacement current boundary occurs where the 'relaxation time' is approximately equal to the timescale of the phenomena[5], which occurs at night at ~ 85 km, where there is typically a sharp ledge in conductivity[3].) This charge initiates a comparable decrease in the cloud's top-to-bottom electric flux which induces an increase in the cloud-to-Earth electric field $\approx \Delta Q/Q_L$.

Triggering and synchronization can be accomplished using the following argument. Electric pulses of P% of the breakdown field, superimposed on a field increasing at R% per second, can synchronously trigger lightning for pulse spacings T for which $P > RT$. Consider a case where R is of the order of 1%s^{-1} and Q_L is taken as a nominal 100 C. Ionospheric charge separation of several coulombs capacitively coupled to the thundercloud would be required for synchronization of lightning-induced whistlers with echoes spaced by several seconds. Electric field measurements made in the ionosphere[6] associated with lightning and whistlers, when combined with appropriate ionospheric conductivities and area coupling to the thundercloud, imply a charge separation of several coulombs, but the authors caution that non-ohmic processes may be involved. It is important to determine the actual ionospheric currents associated with lightning and whistler echoes to assess the suggested mechanism.

A recent computer model of transient ionospheric fields of lightning due to Baginski[5] shows that ~ 10% of the charge lowered to Earth in a cloud-to-Earth lightning stroke is separated in the ionosphere on a timescale of milliseconds after the stroke. (The model, which uses Maxwell's equations, also shows that quasistatic arguments, such as are used here, are valid on this timescale.) As larger lightning strokes can separate tens of coulombs[4], this may be suffient to explain the observation that lightning is sometimes triggered by other lightning at distances up to tens of kilometres. It is noted that such triggering, which may also be associated with solar-activity-related, high-energy electron precipitation events, would tend to increase the proportion of cloud-to-Earth flashes. As cloud-to-Earth flashes contribute directly to the global circuit currents[3], the cumulative effects of both solar-related and subsequently triggered cloud-to-Earth lightning may help explain the increases related to solar/geomagnetic activity in the fair weather field and ionospheric potential reported by numerous workers and summarized by the late D.E. Olson[7].

LESLIE C. HALE
Communications and Space Sciences Laboratory,
The Pennsylvania State University,
University Park, Pennsylvania 16802, USA

1. Armstrong, W.C. *Nature* **327**, 405-407 (1987).
2. Rycroft, M.J. *Nature* **327**, 368-369 (1987).
3. Hale, L.C. in *Weather and Climate Responses to Solar Variations* (ed. McCormac, B.M.) 311-321 (Colorado Associated University Press, Boulder, Colorado, 1983).
4. Uman, M.A. in *Lightning* (Dover, New York, 1984).
5. Hale, L.C. & Baginski, M.E. *Nature* **329**, 814–816 (1987).
6. Kelley, M.C. *et al. J. geophys. Res.* **90**, 9815-9823 (1985).
7. Olson, D.E. in *Weather and Climate Responses to Solar Variations*. (ed. McCormac, B.M.) 483-488 (Colorado Associated University Press, Boulder, Colorado 1983).

Why plants are green

SIR—Goldsworthy[1] ascribes the light absorption properties of the photosynthetic pigments of higher plants to initial competition with prokaryotic organisms that contain bacteriorhodopsin. In this context it is wise to recall the limitations that the chemistry of the pigments places on the plasticity of photosynthetic systems. The green colour of plants could then be argued to be an 'accidental' choice forced on them because of the spectral limitations of those otherwise immensely adaptable components of bioenergetic membranes — the metalloporphyrins.

It would be odd to argue, for example, that the prosthetic group of haemo-

globin, cytochromes or catalase was chosen for its light-absorption properties. Similarly the 'primitive' pigment of an ancestral photosynthesizer need not initially have been selected for optimum absorption of light. Clearly the bioenergetic function of even 'modern' chlorophyll depends as much on its properties as a lipid-soluble electron carrier as on its absorption of light energy[2].

Once a line of organisms opts for a metalloporphyrin structure to carry out any function, however, it is limited by the chemical properties that arise from the high symmetry of the underlying conjugated carbon skeleton. In chemical terms symmetry is expressed as degeneracy in the energy states available to a molecule. Such degeneracies are the starting point for the now classical analysis of metalloporphyrin spectra[3]. This describes how the two-band spectrum of a metalloporphyrin arises: a strongly allowed ultraviolet-blue absorption (labelled B) and a weak green absorption that is formally forbidden (labelled Q). The breakdown of the strict symmetry of the molecule is expressed by Gouterman's parameter v. The square of the dipole moment (r') for the weak green absorption determines the intensity of the absorption according to

$$(r')^2 = (vR-r)^2 \qquad (1)$$

where R and r are the dipole moments associated with the B and Q transitions, respectively, in the perfectly symmetrical molecule and v is a mixing parameter. From this it is clear that a breakdown in symmetry mixes the intensity of the B transition into the weak Q to an extent controlled by v. But this is exacted at a price. The separation in energy between the bands is

$$2A''_{1g}(1+v)^2 \qquad (2)$$

where the factor A''_{1g} is a parameter expressing the fairly constant shift in the energies of excited states by interelectronic repulsion.

In this quantum-mechanical system, then, contrary to 'common sense' the borrowing of intensity from the strong B by the weak visible Q absorption increases as the separation between them increases. The red-shift of a metalloporphyrin visible band is connected with its increased intensity v. To achieve a reasonable visible intensity chlorophyll breaks the metalloporphyrin symmetry by substituting a saturated C–C bond where the conjugation strictly requires partial double-bond character. According to the expressions above this cannot help but shift the Q band from the green into the red — giving a pigment that is green to the eye.

Too great a shift into the red, however, can yield a pigment lacking an excited state that is energetic enough for the photolysis of water; a blind alley now occupied by prokaryotes using bacteriochlorophyll, which contains two conjugation-breaking saturated C–C bonds[4].

It seems plausible, then, that chlorophyll represents a compromise where a metalloporphyrin's visible absorption is maximized without losing the ability for oxygenic photosynthesis so that the unique properties of the metalloporphyrin electronic system as a biological electron 'solvent' can be effectively exploited.

MARTYN SYMMONS
School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

1. Goldsworthy, A. *Nature* **328**, 207–208 (1987).
2. Nicholls, D.G. *Bioenergetics* (Academic, London, 1982).
3. Gouterman, M. *J. chem. Phys.* **30**, 1139–1161 (1959).
4. Olson, J.M. *Science* **168**, 438–446 (1978).

LHRH agonists and human breast cancer cells

SIR—The use of luteinizing hormone releasing hormone (LHRH) agonists to achieve medical castration in patients with metastatic breast and prostate cancer has been widespread in recent years. The primary mechanism by which these compounds block steroid hormone production by the gonads is by inhibition of the release of luteinizing hormone by the pituitary. Miller *et al.*[1] raised the issue of direct inhibitory effects of the LHRH peptides on breast cancer cells. Low-affinity binding activity in malignant but not normal breast tissue has also been reported[2,3]. In studying the effects of a variety of LHRH-related peptides on the growth of MCF-7 hormone-responsive breast cancer cells *in vitro*, however, we have been unable to reproduce the original inhibitory effects[1].

Experiments designed to duplicate the treatment conditions used by Miller[1] did not reveal any inhibitory effects of LHRH or related peptides on MCF-7 cell growth. No change in the doubling time of the MCF-7 cells was caused by any of the compounds studied (see table). There was no effect by Leuprolide (leuprovelin) even when used at a concentration as high as 10^{-6} M. Because of the lack of inhibitory effects of any compound tested, many other treatment schedules and media conditions were tested. They included treatment every day, every other day and every third day with or without concurrent media changes and treatment in media supplemented with 5% charcoal-stripped, sulphatase-treated calf serum with or without oestradiol (10^{-8} M) added to the media. No inhibition was observed.

These observations suggest that clonal variation among MCF-7 cells alters their response to LHRH and its related peptides. In addition, because of the high concentrations of peptides required for inhibition, when it is observed, and the low affinity ($K_d = 10^{-6}$ M) of LHRH binding[2,3], it is unlikely that LHRH agonists exert significant direct biological effects on breast cancer cells *in vivo*. Perhaps the observations of Miller *et al.*[1] and Eidne and co-workers[2,3] are best viewed as insights into the possibility of an autocrine regulatory function for peptide hormones in mammary cells.

Doubling times of cells

Treatment	Doubling time (% control)
Control	100
LHRH (10^{-7} M) (Sigma)	93
Leuprolide (10^{-6} M) (Tap)	93
(10^{-7} M)	100
(10^{-9} M)	100
(10^{-12} M)	110
Buserelin (10^{-7} M) (Hoechst)	100
L2636 (10^{-7} M) (Sigma)	100
L4261 (10^{-7} M) (Sigma)	107

Cells were grown in Dulbecco's minimal essential medium supplemented with 10% fetal bovine serum (FBS), gentamicin (40 mg l^{-1}) and glutamine (0.6 g l^{-1}). Medium was changed daily and treatment substance added daily. Cells were counted every other day by cell counter (Particle Data). No insulin was added to the media. Control cell doubling times ranged from 24 to 33 h in the presence of FBS and phenol red. Time points for each treatment were performed in triplicate.

G. WILDING
M. CHEN
E.P. GELMANN
Medicine Branch, National Cancer Institute, Bethesda, Maryland 20892, USA

1. Miller, W.R., Scott, W.N., Morris, R., Fraser, H.M. & Sharpe, R.M. *Nature* **313**, 231 (1985).
2. Eidne, K.A., Flanagan, C.A. & Millar, R.P. *Science* **229**, 989 (1985).
3. Eidne, K.A., Flanagan, C.A., Harris, N.S. & Millar, R.P. *J. clin. Endocr. Metab.* **64**, 425 (1987).

MILLER REPLIES — Several groups have contacted me indicating an inability to reproduce our results in their laboratories and I think that, as Wilding *et al.* suggest, clonal variation of cells and culture conditions can in part explain these discrepancies. Certainly, we have in our possession other MCF-7 cells which do not respond in the same way to LHRH agonist and we have reported that, in the presence of insulin, it is difficult to show effects of LHRH agonists (W.R. Miller *et al.*, *Monogr. Ser. EORTC*, Vol. 18). An exchange of cells with Wilding is in progress to resolve these issues.

Nevertheless, there are other reports of direct effects of LHRH analogues on breast cancer cells. Thus Blankenstein and co-workers (*Eur. J. Cancer clin. Oncol.* **21**, 1493; 1985) showed inhibitory effects on MCF-7 cells; Wiznitzer and Benz (*Proc. Ann. Am. Ass. Cancer Res.* **25**, 208; 1984) inhibitory effects on T-47D cells and, as Wilding *et al.* point out, Eidne *et al.* reported inhibitory effects of LHRH antagonists on several cell lines.

W.R. MILLER
University Department of Clinical Surgery, Royal Infirmary, Edinburgh EH3 9YW, UK

Of the utmost gravity

John D. Barrow

300 Years of Gravitation. Edited by S. W. Hawking and W. Israel. *Cambridge University Press: 1987. Pp. 684. £45, $69.50.*

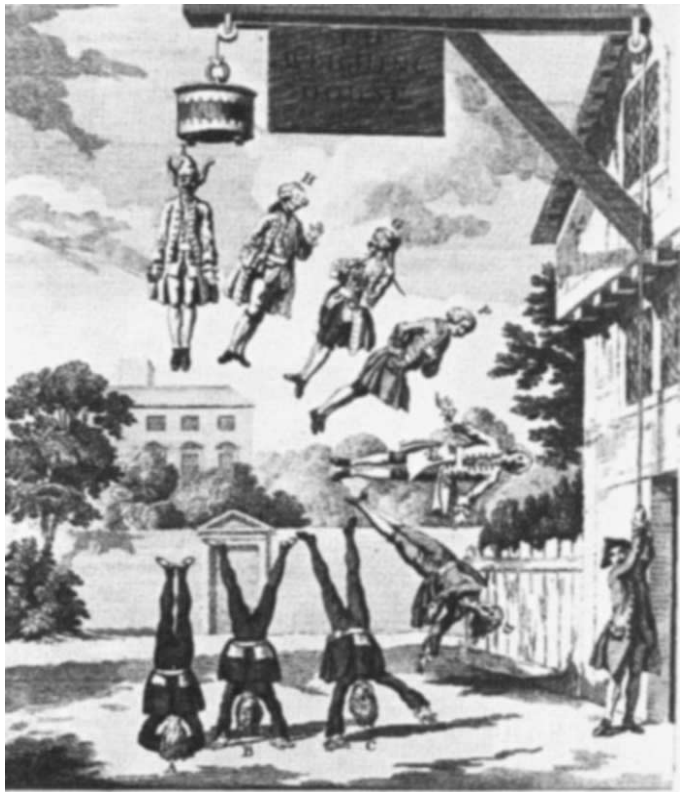
IT all began one August day in 1684 when Dr Edmond Halley paid a visit to Isaac Newton's rooms in Cambridge; of that meeting Abraham de Moivre tells us that

> After they had been some time together, the Dr asked him what he thought the Curve would be that would be described by the Planets supposing the force of attraction towards the Sun to be reciprocal to the square of their distance from it. Sr Isaac replied immediately that it would be an Ellipsis, the Doctor struck with joy & amazement asked him how he knew it, Why saith he I have calculated it, Whereupon Dr Halley asked him for his calculations without any further delay, Sr Isaac looked among his papers but could not find it, but he promised him to renew it, & then send it to him.

Three months later, true to his word, Newton sent a document to Halley that made good his remarkable claim. Encouraged further by Halley it took Newton but 14 months to transform and extend this manuscript into another entitled *Philosophiae Naturalis Principia Mathematica*. It was destined to become Book I of the *Principia*. Two further parts followed in March and April 1687 just two and a half years after Halley's first fortuitous visit.

The first edition of the monumental *Principia*, comprised of these three manuscripts, was published in London on 5 July 1687 under the imprimatur of Samuel Pepys, the President of the Royal Society (whose diary-keeping days were by then far behind him). A second edition appeared in 1713, and the third (and last) in 1726. Three years later the only full English translation ever made of the original Latin text was completed by Andrew Motte and published as a two-volume set.

It is this chain of events that inspired the appearance of this tercentenary volume as a celebration of the most impressive and comprehensive work of science ever written. Today, no practising scientist reads the *Principia* out of necessity but it is the foundation stone upon which the entire edifice of mathematical physics and astronomy rests. Newton's laws of motion and gravitation, although superseded by Einstein's, provide an approximation to them that hold good in all the situations we encounter in everyday life. More important than the formulation of these rules by which Nature moves was the thoroughly rigorous and 'modern' approach that Newton introduced. The unity of Nature was displayed by his ability to derive far-reaching and diverse results about the extraterrestrial physical world from a bundle of innocuous-sounding assumptions that stemmed from observations of mundane local objects such as apples and prisms.



Weighty matters — the condition of man from A, absolute gravity through E, horizontal or good sense and F, wit to I, absolute levity or stark fool. A contemporary engraving from the collection at the British Museum.

Despite its vast scope, to most scientists the *Principia* is synonymous with gravity and with it the first statement of a universal law of gravitation. This fact is reflected in the title and contents of Hawking and Israel's collection of 14 articles. With only one or two exceptions, the individual contributions are of the highest quality and expound their subjects in impressive style with careful attention to didactic completeness and recent historical development.

Before dwelling on the highnotes of the individual articles, however, I must first confess I found this volume disappointing in some respects. Unfortunately, the connection with Newton's work is not well-explored — nine of the articles manage to avoid any mention of Newton or the *Principia*, and one author has substituted a brief essay written for another occasion beneath a note that his real contribution was not ready in time. But most striking is the narrow scope of the articles commissioned. They are confined, with the exception of Israel's meticulously researched history of white dwarfs, neutron stars and black holes, to current research frontiers in general relativity. Where are the articles on modern dynamics? Celestial mechanics? The N-body problem? Galactic dynamics? The stability of the Solar System? The exploration of the Solar System's outer planets and the mysterious dynamics of their ring systems? All of these subjects are just as interesting, just as 'gravitational' and considerably more relevant to the heritage of the *Principia* than some subjects that have been included (and, in the case of the inflationary universe, duplicated). And was there really not enough space for a real historical article on gravitation by a professional historian?

Although the book is not an entirely satisfactory celebration of the *Principia* and the last 300 years of gravitation, it does provide an outstanding survey of the last decade of gravitation which every astrophysicist should try to read from cover to cover. The only blemish is the publisher's error in not providing an index, an economy that is made harder to appreciate by the fact that they have been carried away by the centenary implications to such an extent that the price of this book is precisely one hundred times that of the first edition of the *Principia*.

The articles fall into four groups. The first — Cook on gravitation experiments, Damour on equations of motion, Will on tests of gravitation theories and Thorne on gravitational waves — covers subject areas with long pedigrees which are firmly anchored to experiment and observations. Each of these authors is a master at organizing and making sense of a bewildering array of material, and each has produced a

beautiful overview of his subject.

In the second group are articles of an astrophysical nature by Blandford and Rees. They deal with complicated phenomena that are only partially understood and not so well suited to systematic exposition. Faced with a difficult and wide-ranging task, both authors display not only the important facts and theories of high-energy astrophysics and cosmology, but also something of the special talents that are required to work successfully in these fields where all manner of different techniques — mathematical, statistical, physical, observational, speculative — are required to produce explanations of phenomena where many different physical effects intermingle to confuse us.

In the third group are articles of a more speculative variety on the frontiers of theory: Penrose describes the paradoxes of quantum reality and the possible role of gravity in the collapse of the wave function and the determination of the arrow of time; Linde, and Guth and Blau discuss the inflationary universe which the first two authors have done so much to develop and promote; Vilenkin manages to tell you all you need to know about cosmic strings in a 23-page article that is a model of clarity; Schwarz supplies a comprehensive yet readable contribution on superstrings and where he hopes they are going (only time will tell whether they get there via Stockholm); Hawking gives an update of his programme to calculate and interpret the wave function of the Universe, and explains his latest thinking about the arrow of time and the necessity for the weak anthropic principle to be employed in quantum cosmology; and Crnkovic and Witten provide a brief and somewhat technical exposition on canonical formalism. All the articles in this group deal with theoretical subjects that are as yet devoid of experimental test or decisive observational contact.

Finally, in the last group are an essay on newtonianism by Weinberg and the fascinating historical article by Israel. This charts the thinking and writings of astronomers on 'dark stars' from the eighteenth century to the present, and includes interesting accounts of his own thoughts at crucial stages in the interpretation of mathematical results about black hole mechanics and thermodynamics during the 1960s and early 1970s.

One of the strongest impressions to be gained from these articles is of the diversity of work in modern gravitation physics and the way in which gravitation is at last ceasing to be the aloof geometrical cousin of the rest of physics. In astrophysics it merges with the complexities of high-energy radiation processes to create spectacular physical phenomena. The search for gravitational waves exploits the latest technologies to provide astronomers with the promise of a new vista on the Universe. And in the inner space-time of elementary particles there are the first glimmerings of a convincing theory which might conjoin gravity with the other forces of Nature.

Who can doubt that Newton would be pleased but not surprised at these advances towards a unified description of the physical world and all its forces. For we find that 300 years ago, in the Preface to his stupendous book, he could write that, having determined "the motions of the planets, the comets, the moon and the sea", he was unfortunately unable to determine the remaining structure of the world from the same propositions because

> I suspect that they may all depend upon certain forces by which the particles of bodies, by some causes hitherto unknown, are either mutually impelled towards one another, and cohere in regular figures, or are repelled and recede from one another. These forces being unknown, philosophers have hitherto attempted the search of Nature in vain; but I hope the principles here laid down will afford some light either to this or some truer method of philosophy.

How right he was. □

John D. Barrow is a Lecturer at the Astronomy Centre, University of Sussex, Brighton BN1 9QH, UK.

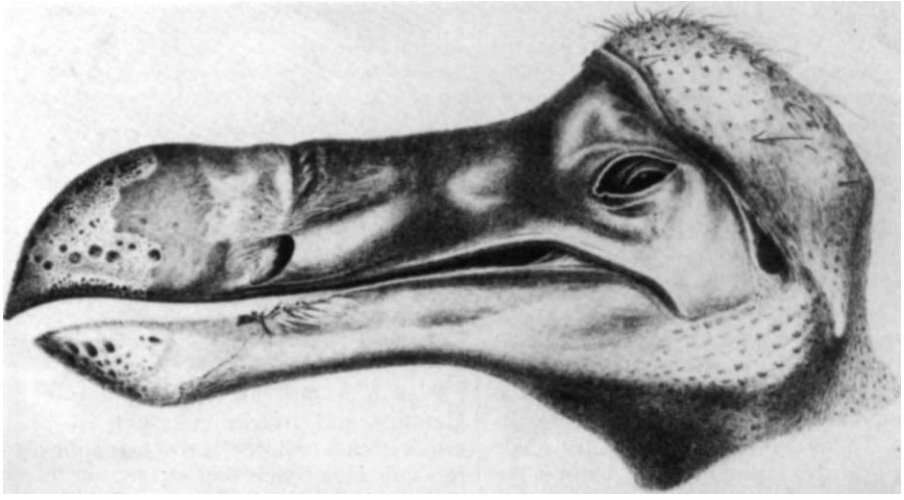
Skimming through space

James Trefil

Space 2000: Meeting the Challenge of a New Era. By Harry L. Shipman. *Plenum: 1987. Pp.431. $19.95.*

In this book, University of Delaware astronomer Harry Shipman undertakes nothing less than a survey of space science and technology, both past and future, along with an introduction to the basic science (primarily astronomical) that is being done above the Earth's atmosphere. Despite the book's 400-plus pages, the sheer volume of the subject matter dictates that the coverage must of necessity be of the "broad and shallow" as opposed to the "deep and narrow" variety. Thus, the book is not intended for those already familiar with the space programme, but as an introduction for those who are not.

The book is divided into three sections. The first concerns the basic physics of satellite motion, the history of the space programme and an analysis of space technology. As you might expect from an author who has been deeply involved in the programme and was, in addition, the public relations and education officer of the American Astronomical Society, this part is very well done. Shipman's extensive experience with the space programme comes through very clearly, and there are many useful insights. To cite one among many, he points out a fundamental difference between the proposed Space Station and the Shuttle — an obvious distinction often overlooked by the Station's critics: the Space Station is modular, which means that unlike the Shuttle it can be built in stages; building it will not, therefore, necessarily drain money from other programmes, as the all-or-nothing Shuttle programme did in the 1970s. I



Gone but not forgotten — the dodo head preserved in the Ashmolean Museum, Oxford, as depicted in H.E. Strickland and A.G. Melville's The Dodo and Its Kindred *of 1848; the rest of the specimen (the right foot apart) was thrown out in 1755, by which time it could not be replaced. The picture is reproduced from* Extinct Birds *by Errol Fuller, published today in Britain by Viking/Rainbird, price £20. In the United States the book will be published by Facts on File.*

suspect this section of the book will provide little nuggets like this for most readers who, like me, have followed the space programme in an offhand way, but never thought about it very deeply. This section also provides a useful (and sobering) update on various space industry enthusiams of the late 1970s. Solar power satellites (a victim of the oil glut), the purification of drugs in space (now facing competition from the new technology of genetic engineering) and microgravity manufacturing are all dealt with in a realistic way. The bottom line is that it is only communications technology that seems truly to be flourishing in space.

In the second section of the book — by far the longest — Shipman attempts a survey of all the science that can be learned from space. Here the broad nature of the subject matter causes problems. There simply isn't enough attention devoted to any single topic to cover it adequately, and the result is an unsatisfactory potpourri of planetary science and astrophysics, with snatches of other subject matter thrown in for good measure. The brevity of the discussion of complex points will make it very difficult for someone not already familiar with a subject to understand what is going on. To take just one example, the development of life on Earth is passed over in just three pages — hardly enough to instruct the neophyte, but too much for the *cognoscenti*.

In the final section Shipman returns to his strong suit with a discussion of the future of the space programme. He presents a well thought out survey of the role of man in space, mercifully free from the kind of either/or rhetoric that has characterized the debate between NASA and some members of the space science community. For Shipman, the question is not whether humans will be involved in future space missions, but where the contributions of human beings make the extra cost of providing a livable environment worthwhile. He points out that it was only after an oceanographer flew on a Shuttle mission in 1984 that it was appreciated that detailed eddy and current structures in the oceans could be seen from orbit. Finally, in a chapter frankly labelled as speculation, the "far out" aspects of the space programme are mentioned. Moon bases, communications and materials technology, and the various strategies for a Mars mission make a fitting ending to the book.

The canvas of *Space 2000* is painted with a broad brush, taking in technological, scientific and political issues, past and future. The book has all the strengths and weaknesses of that approach. □

James Trefil is Clarence Robinson Professor of Physics at George Mason University, 4400 University Drive, Fairfax, Virginia 22030, USA. His latest book is Meditations at Sunset *(Scribners', 1987).*

Living in the past

J.N. Graham Ritchie

The Stonehenge People. By Aubrey Burl. *Dent: 1987. Pp.249. £16.*
The Stonehenge People: An Exploration of Life in Neolithic Britain 4700–2000 BC. By Rodney Castleden. *Routledge & Kegan Paul/Methuen: 1987. Pp.282. £14.95, $25.*

THE message of these volumes is of an overworked seam of archaeological writing. Surely not all popular books about prehistoric Britain need have Stonehenge in the title, nor the added inducement that they are describing 'real' people rather than the potsherds and postholes of the archaeological evidence?

Of the two, the Burl volume is less populist. After setting the scene amid the earthen long barrows and causewayed enclosures of the Wessex downland, Burl uses the sequence of construction and reconstruction at Stonehenge as the basis for a wide-ranging discussion of prehistoric society from around 3200 BC into the first millennium BC. By concentrating primarily, though by no means exclusively, on the Stonehenge region, Burl places in perspective the surviving artefacts, monuments and graves, and creates a series of vignettes of downland society and the wider geographical contacts on which it drew. His valuable interweaving of the historical aspects of the archaeological activities in the area, from the Duke of Buckingham in 1620, who appears to have dug into the centre of Stonehenge, helps to place our own attitudes into a wider context.

Burl is particularly level-headed about the possibilities of astronomical observation in prehistoric times. On the one hand he stresses the likely influence that movements of the celestial bodies may have had within non-literate societies, on the other he affirms that "Stonehenge was never used for scientific observations. It was never inhabited by astronomer–priests searching for explanations of eclipses or detecting tiny lunar oscillations". In general there is too much detail and insufficient illustration, but I much enjoyed those sections where Burl recreates the bustle and organization during the construction of Stonehenge, with assessments of such time-consuming operations as the dressing of the sarsen blocks.

Castleden's volume is more general and, although covering the same chronological span, uses broader brush-strokes and fewer detailed examples. The author is a geographer and geomorphologist by training, and among the most interesting sections are those on crops and livestock and flint mines. But overall he has successfully compressed a great deal of information into the text and illustrations,

From Castleden's *The Stonehenge People*

Open door — Trethevy Quoit in Cornwall, a portal dolmen which was the entrance to an earth burial mound.

making a readable and useful book. Where a new or alternative idea is presented, it is clearly described; not all need be accepted, however attractive. The idea that there was large-scale trading of pottery vessels between Orkney and Wessex (in the style known as Grooved Ware) in the third millennium BC strains the imagination, though I am glad to see the early date of the Orcadian examples underlined.

Attempts to conjure up prehistoric societies in book form, as opposed to museum display or on-site explanation, have to strike a balance between a textual description of the scattered archaeological evidence and a graphic illustration of the material, so that the reader will readily grasp what is meant by a Grooved Ware vessel, mace-head, henge-monument and the like. Part of the appeal of the past is a visual one. In this respect Castleden's book is more successful, with some unusual and arresting images, although many of the photographs are far too dark.

Both volumes are well balanced in approach and style, taking the reader as far as the evidence is likely to allow and no further. Both seem to be designed for the book-club market and a large number of copies will perhaps be bought unseen. I suspect that the general reader will enjoy handling and reading the Castleden book more, though archaeologists will be disconcerted by such sentences as "Prehistorians have understandably been swept along by the ritual preoccupations of the neolithic people . . .". □

J.N. Graham Ritchie is an Archaeological Investigator with the Royal Commission on the Ancient and Historical Monuments of Scotland, 54 Melville Street, Edinburgh EH3 7HF, UK.

Where's the action in endocrinology?

J. R. Tata

Hormones. By Anthony W. Norman and Gerald Litwack. *Academic: 1987. Pp.806. $49.95, £37.*

HORMONES are big business for publishers, but useful, unconventional textbooks of modern endocrinology are few and far between. The hopes for something new in Norman and Litwack's *Hormones* are partially fulfilled. According to the authors, the book "provides a comprehensive treatment of human hormones" for first-year medical students, graduate students and advanced undergraduates in the biological sciences. Only parts of it can be said to be truly useful for this group of readers, but the book does succeed in fulfilling the authors' claim to "filling a void" in its emphasis on the cellular and molecular aspects of the subject.

Except for chapters on "general considerations", steroid hormones and growth factors, the 16 chapters constituting the bulk of the book deal with hormones according to which gland makes them. This is the traditional approach and it results in considerable repetition of topics such as mechanism of action of hormones, receptors and multi-hormone systems. Had the authors divided much of the material into individual accounts of general issues — hormones and second messengers, nuclear receptors and gene structure and function, membrane receptors and phosphorylation, co-evolution of peptide hormones and their receptors, feedback mechanisms and the integration of neural and endocrine systems, for example — not only would they have avoided such repetitiveness but they would have provided a much-needed consideration of hormones as part of a continuously evolving system of biological regulation via chemical signals.

For first-year medical students, the descriptions of anatomy of endocrine glands, physiological actions of hormones and clinical manifestations of endocrine disorders will certainly be helpful. But one may well ask of what value are the numerous diagrams of protein and DNA sequences of peptide hormones, receptors and their genes for students of medicine or general biology? Norman and Litwack say that this is to impress the reader of the recent rapid progress in molecular biology, particularly in recombinant DNA and gene cloning. Couldn't this have been said in a few sentences? Although giving the book a cachet of trendiness, the exaggerated emphasis on these new technologies makes it less definitive and more likely to date quickly. Quite rightly, the authors have stressed the cellular and molecular basis of hormone action, but they fail to emphasize that this is simply a reflection of what we know at a given time about biochemical, cellular and molecular mechanisms of biological regulation as a whole.

The material is well presented with clear illustrations (the anatomical diagrams of endocrine glands and the portrayals of biosynthetic and metabolic pathways of hormones are particularly good). Many of the models of hormone action are reminiscent of those in *Biochemical Actions of Hormones,* the highly successful series edited by Litwack; some, though, need revision. I could spot only few errors such as cyclic AMP being described as a hormone on p.40. The bibliography, listing review articles and research papers separately, is remarkably up to date, although one frequently notices omissions of references to pioneering work. There is also an informative appendix at the end, with a compilation of the sites of synthesis and actions of the major mammalian hormones, their concentration in blood and diseases associated with them, but the inclusion of a list of Nobel laureates is puzzling.

All things considered, Norman and Litwack's *Hormones* will find an appreciative audience of first-year medical students and undergraduates reading biology. It will also serve as a good source book for cellular and molecular endocrinologists — at least for the next five to ten years. □

J.R. Tata is Head of the Laboratory of Developmental Biochemistry at the Medical Research Council National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.

Enigmatic reprise

Alan Nunn May

Klaus Fuchs, Atom Spy. By Robert Chadwell Williams. *Harvard University Press: 1987. Pp.267. $25. To be published in Britain on 23 November, £19.95.*

ANOTHER biography of Klaus Fuchs so soon after Norman Moss's (reviewed here on 4 June*) may seem too much of a good thing, but there are interesting contrasts in the two treatments. As a journalist, Norman Moss based his account largely on interviews. Professor Williams, a historian, uses mainly documents, among them the papers of Chadwick, Cockcroft, Peierls, Cherwell, the memoirs of Jurgen Kuczinski (Fuchs's first contact) and the FBI papers on the case, which include MI5 documents which are still secret in Britain. For the first time we can read Fuchs's full technical confession, and Harry Gold's confession to the FBI.

All this yields many new insights. But at times Professor Williams is tempted rather far from his main topic, and he does not always achieve complete accuracy. To give three examples of errors among many, he says that Britain and the Soviet Union were still enemies during the first six months of 1941; that Max Born's laboratory in Göttingen was one of the great centres of nuclear physics; and that the cyclotron used alternating magnetic fields to accelerate particles.

On the other hand, Professor Williams's explanation of the English law of treason, and why Fuchs was not charged with this crime, but only with the offences under the Official Secrets Act, clarifies a topic much confused by previous writers, including Norman Moss. He is also the first to mention the secret Anglo-Russian agreement, promising full exchange of information on all new weapons, which has been glossed over or ignored by atomic historians. But the suggestion that therefore Fuchs was only helping the British to fulfil their treaty obligations would probably have failed to impress the judge at his trial.

About the trial Professor Williams holds strong opinions. He says that there should have been a jury, with witnesses and cross-examination. He holds that the British prevented this to avoid some embarrassing disclosures, such as Fuchs's role in the secret British bomb project; the Russian penetration of MI5 and MI6; the (official) sharing of intelligence with the Russians during the war, and contacts between Jurgen Kuczinski and cabinet ministers. The first of these points might have had some influence, but the others were either unknown at the time, or were barely relevant to the Fuchs case. One must exonerate the British of perverting justice. There was no jury because a plea of guilty leaves nothing for a jury to do.

It is a pity that Norman Moss and Professor Williams did not collaborate and produce a joint work. Moss's personal contacts, Professor Williams's documentation and a mutual correction of errors could have produced the definitive biography. □

Alan Nunn May was a member of the British–Canadian atomic project in Montreal from 1943 to 1945. In 1946 he was convicted of passing information to the Soviet Union, and was sentenced to ten years in prison. In 1961 he was appointed Professor of Physics in the University of Ghana, and retired in 1978.

*Moss's book *(Klaus Fuchs: the Man Who Stole the Atom Bomb),* which was published earlier this year in Britain by Grafton, becomes available in the United States on 2 November. Publisher is St Martin's, price is $15.95.

Computing the best for Europe

Dennis Jennings

Computer networks are developing fast, but Europe may be heading down the wrong road.

THE pace and complexity of modern research has dramatically increased the need for fast, reliable and easy to use communication. Developments in communications and information processing have made it much easier for the many thousands of researchers in academic and government institutions to communicate with each other and with industry, and to undertake collaborative work. Researchers now depend on computing and communications tools and are handicapped without them.

In Europe, the growing emphasis on collaboration through programmes such as ESPRIT and RARE is increasing the demand for closer international co-operation and improved communications. In line with the development of such programmes, there needs to be a parallel development of the European communications infrastructure for research. To understand the issues this raises, some background information is needed.

The European Academic and Research Network (EARN) is a computer network linking tens of thousands of academics in Europe to their peers and colleagues throughout the world. The network was started in 1983, and was modelled on BITNET in the United States. The model is of a store-and-forward network, where mainframe computers are linked together by leased lines and, using standard software, provide electronic mail and file transfer services to users. The BITNET model is simple but elegant: new institutions connect to the network by leasing a telephone line to an existing member and the network tables are updated.

EARN is not quite the same as BITNET, in that each country has a key or 'backbone' computer site, and other institutions in each country are connected to these backbone sites. The national backbone sites are themselves connected together by international lines. The basis of EARN, then, is a European backbone network, linking to and incorporating the component national networks.

Since 1983 EARN has grown rapidly, and there are now 20 countries connected to the network (Fig. 1), including the Ivory Coast in Africa, and approximately 600 host computers. Most of these computers are either DEC VAX/VMS or IBM VM/CMS systems. The various national parts of EARN work in several ways — in some countries the network is an extension of EARN; in Britain there is at present a single EARN node computer which acts as a gateway to a separate national academic network; and in other countries a mixture of the two approaches is taken. In addition, EARN is connected to BITNET, and thus to ARPANET, NSFnet, CSNET and to most of the other academic networks in the world. Upwards of 30,000 academics now use EARN for their day-to-day research activities.

EARN provides three basic services — file transfer, electronic mail and remote job execution. On top of these, a variety of additional services are available, including electronic mailing lists, file servers and database services accessed by electronic mail.

The EARN services are heavily used, most particularly electronic mail between individuals or groups of researchers who can discuss and plan new work, prepare a paper, submit a new proposal or complete a report. Researchers may also subscribe to one or more of the many electronic mailing lists on topics bearing on their research, and in this way they can follow discussions between colleagues throughout Europe and the world and themselves contribute to the debate. Initially, most of these mailing lists were hosted on the US ARPANET network, but a growing number of lists on topics of particular interest to European researchers are hosted in Europe.

The other EARN services are used in a variety of ways — to exchange data; to submit work to larger computers; to transmit experimental data to collaborators at other institutions; to send files from a file server maintained at a particular site. Traffic levels are growing steadily, and many sections of the network are becoming heavily loaded at times.

The EARN organization is modelled on the network itself. EARN is an association, registered in France, and is run by a board of directors, one from each of the countries connected to the network. Each institution with computers connected to EARN is a member of EARN, and the EARN director in each country represents the members. Every two years, or as posts become vacant, the board elects an executive committee to run the association. The day-to-day affairs of the association are run by a management team, currently with offices in Dublin and Paris.

When EARN was started, all the funding was provided by IBM Europe, and by IBM in the various European countries. This included money for the equipment at many of the backbone sites, and for support personnel from IBM. As the network developed, the additional costs of adding new members and running gateways to

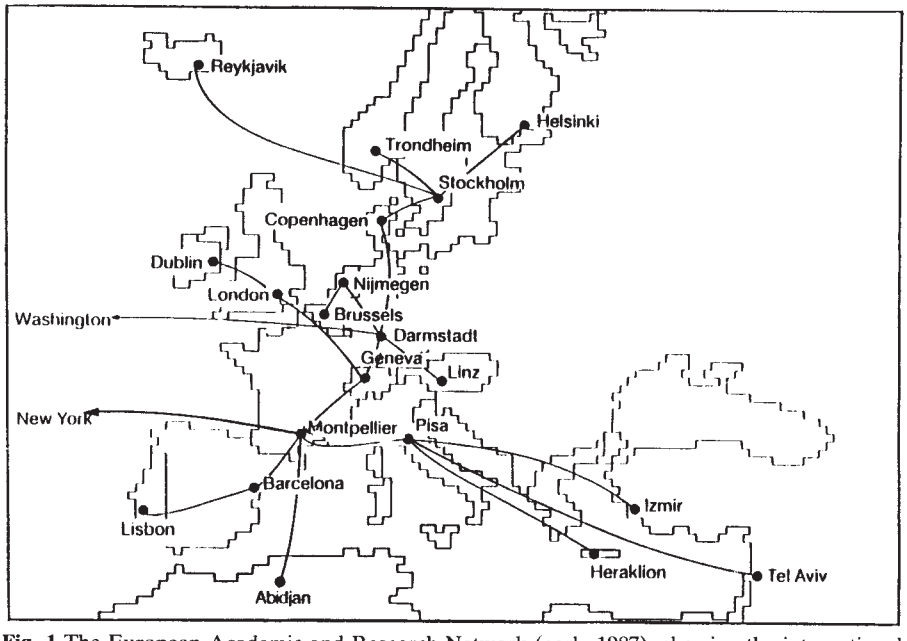


Fig. 1 The European Academic and Research Network (early 1987), showing the international leased lines connecting the EARN backbone sites in each country.

national networks were met by the members in each country. All EARN services are provided free to researchers using EARN.

IBM's initial financial support was for a period of four years, and, after a final grant to the association for the continuing management of EARN, all support from IBM Europe ceases at the end of 1987. Despite predictions to the contrary, this has not proved as difficult a problem as many observers expected. In 1988, EARN will be funded by the members in each of the participant countries. These funds will cover the costs of the EARN backbone, and of the connections between EARN and BITNET. The total amount involved is approximately 1.2 million ECU (European Currency Units; 1ECU=*c*.£0.70). In addition, each national director has arranged local funding to cover the costs of EARN in that country. An important consequence of these financial arrangements is that the costs of EARN are fixed in advance, and will be borne by the member institutions. Thus all EARN services will continue to be provided free.

There were several problems to be faced in the initial development of EARN. The main one was with the national PTTs and with the European Conference of Postal and Telecommunications Administrations (CEPT). Technically, from the point of view of the PTTs, EARN connects different organizations together, and thus each EARN computer is switching third-party traffic. The PTT response to this, through CEPT, was to suggest that all such third-party traffic would have to incur a volume-related charge, on top of the basic leased-line rentals. Further, CEPT insisted that any decision to permit EARN to operate would only be given provided EARN agreed to migrate to the use of the International Standards Organization (ISO) Open Systems Interconnection (OSI) protocols by the end of 1987. For a while the installation of many of the leased line circuits was delayed. Indeed, at one stage, it was only possible for academics in northern Europe to send messages to southern Europe via North America and the connections to BITNET. Eventually, however, these problems were sorted out. When it became clear that EARN was an independent academic network, and not IBM dominated, as some of the PTTs apparently had feared, and when EARN's commitment to the eventual migration to the use of ISO/OSI standards, as these become available and practical, was understood, most of the initial difficulties disappeared. Some PTTs deferred the issues. Some decided that EARN should be permitted to operate. Others decided that the idea of volume charging on leased lines should be dropped. By mid-1985, all EARN links were fully operational.

The migration of EARN to the use of ISO/OSI standards remains a key issue. A draft migration strategy is being developed, and it is hoped that it will be adopted by the board later this year. This strategy is not yet finalized, but will probably contain two separate approaches: bottom-up, involving the replacement of the current low-level protocols by an X.25-packet-switched network to handle the basic switching of traffic; and top-down, to replace the current EARN application protocols by the OSI applications. However, it is becoming clear that the implementations of the ISO/OSI protocols are immature, and do not provide the same level of service that EARN users currently enjoy. This is expected to change quite rapidly over the next two or three years. In addition, there are technical issues which affect the actual migration process. Nonetheless the EARN board remains committed to a migration to ISO/OSI standards as soon as practicable. In this matter, EARN will be working closely with RARE.

Table 1 Membership of RARE

National members	*National network*
Austria	UNA
Belgium	ABUT/BVT
Denmark	—
Finland	FUNET
Germany	DFN
France	REUNIR
Greece	ARIADNE
Iceland	SURIS
Ireland	HEANET
Luxembourg	—
Netherlands	SURFNET
Norway	RUNIT
Portugal	—
Spain	IRIS
Sweden	SUNET
Switzerland	SWITCH
United Kingdom	JANET
Associate members	
Hungary	
Republic of Korea	
International members	
CERN	
EARN	
ECFA	
NORDUNET	

RARE (Réseaux Associés pour la Recherche Européene; Associated Networks for European Research) is a European organization of national research networks, established in 1985, which currently has full national members in 17 countries (Table 1). RARE also includes a number of countries interested in full membership, a number of associate (non-European) national network members, and several international members, including EARN.

The objective of RARE is to establish a communications infrastructure for academic research in Europe. However, unlike EARN, RARE plans to achieve this, not by building a European networking infrastructure, but by unifying, or harmonizing, the national research networks and connecting them through the use of existing public telecommunications facilities, specifically public X.25 PTT services. To achieve this, RARE supports the principles of Open Systems Interconnection (OSI) as defined by the ISO, and is working to establish functional standards which specify the particular choices of ISO/OSI standards which actually make interworking using OSI possible. Thus RARE fully supports the European standards organization CEN/CENELEC to establish these functional standards. In this, and its view of European networking, RARE is actively supported by a Directorate (DGXIII) of the Commission of the European Communities (CEC). RARE's activities are pursued through eight Technical Working Groups, and the annual Networkshop.

RARE is an association legally constituted in the Netherlands, and is run by a council of administration made up of delegates from each national member, and observers from other members. It is funded through annual subscriptions paid by the member countries. These subscriptions range from 1,000 ECU for associate and international members, through 2,000 ECU for the smallest national member, to 20,000 ECU for the largest European countries. In addition, RARE is supported, in terms of participation and organizational support for meetings, by the CEC, and there is hope that the CEC may fund some pilot projects organized by RARE for new services based on ISO/OSI application protocols — the first such project is the X.400 electronic mail project based on the EAN software developed in Canada, and the Gypsy software developed in France. It should be noted that the subscriptions paid by member countries are to fund RARE, not for the network services arising from RARE's harmonization activities. The RARE model for services is that all services will be charged on a usage basis, and these usage charges will be billed directly to each researcher.

So RARE, supported by DGXIII of the CEC, takes the multi-national approach to European networking; that is, the view that a European academic computer networking infrastructure can be created by ensuring that the various national academic networks obey the same ISO/OSI functional standards in the future, and thus, using the public data network interconnections between the national PTTs, will be able to intercommunicate.

An important activity of RARE is the work undertaken under contract for COSINE. This began in 1986 when RARE was invited to draft an outline plan for the infrastructure of the European Research Network project COSINE. This

Table 2 European network activities

Network	*Types of links*	*US link*
EARN (European Academic and Research Network)	Leased lines	To BITNET from Darmstadt and Montpellier
HEPNET (High Energy Physics Network)	Leased lines and PTT X.25 services	To HEPNET from CERN
SPAN (Space Physics Application Network)	Leased lines and PTT X.25 services	To SPAN from Darmstadt
EUNET (European Unix Network)	X.25, telephone link-up	To UUCP from Amsterdam

COSINE infrastructure is to provide the rapid establishment of an environment to make data communications services available to users from academic and industrial research organizations in Europe, using common carrier services and commercially available hardware and software communications products. The COSINE project is supported by the European governments, and by DGXIII of the CEC, and a total of 800,000 ECU has been provided by the national governments for the one-year specification phase. Like RARE, COSINE seeks to establish a European infrastructure based on a multinational approach, and the COSINE project specifically states that "the implementation phase will involve a shift of emphasis from international and centralized activities within the project to national and decentralized activities". The specification phase, due to end in mid-1988, covers functional standards, charging and accounting, address and route management, directories, testing and diagnostics, PTT facilities, and organization. A secretariat is provided by DGXIII of the CEC.

To complete the picture of academic computer networking in Europe, mention must be made of the main discipline-orientated European-wide networks and of the national academic networking activities (see Table 2).

Although there are many similarities between the activities of EARN and RARE, there are also striking differences. The first is the overall perspective of the two organizations.

EARN is dedicated to providing services to users today; to changing those services to ISO/OSI when this is the proven path to better (or at least equivalent) services; to a market-driven approach to the supply of basic telecommunications services; and perhaps most fundamentally, to a wholly European approach to the provision of computer networking in research, and not a multi-national approach.

RARE and COSINE, on the other hand, are dedicated to enabling the provision of services to users based on the emerging implementations of standard ISO/OSI services; to a standards-driven and PTT-based approach to the supply of basic telecommunications services; and, most fundamentally, to a multi-national view of European networking. In this, it is strongly supported by the CEC, and by the national networking organizations.

The multi-national perspective of RARE, the national networking organizations and the CEC, is, I believe, mistaken, and bodes ill for the future of a truly European development in this, and indeed other, fields. The importance of the European, as opposed to the multinational, approach cannot be underestimated. The European approach is more than just the sum of the component national parts; it creates a new and substantially larger entity which provides strength to Europe as a whole.

The second key difference between the EARN and RARE approaches is in the way the costs of networking services are presented to individual researchers.

EARN is dedicated to providing a European networking infrastructure where both the capacity and the costs of the network are known in advance and fixed. Thus the costs can be borne by the EARN member institution, and are easily budgeted for in advance. As a consequence, EARN services are provided at no charge to the individual researcher, and, within the overall capacity of the network, researchers are encouraged to exploit the network to best meet their individual needs.

RARE and COSINE, by contrast, are dedicated to the use of the PTT data services, where the charges are directly related to usage. Costs are not known in advance, are not fixed and are, in principle, unlimited. Member institutions cannot budget for them, and network services are charged directly to each individual researcher.

In my view, RARE's approach to charging for services will inhibit the development of computer networking to meet the needs of European researchers. A further restriction, if the RARE model is adopted, is that academic computer networking can only be based on those data services for which a general market has been established by the PTTs. Thus, unlike their counterparts in the United States and elsewhere, who can use advanced high-performance networking services as they are developed, European researchers are condemned to wait until new services have become stable and profitable enough to be introduced by all the national PTTs.

One of the first and most important questions to be asked about the development of a networking infrastructure concerns the cost. One should be able to go to the telecommunications supply market and get competitive quotations for the services required, and, indeed, get a variety of technical responses with perhaps widely varying costs and performance characteristics. One should then be able to choose from these offerings to build the appropriate and most cost-effective solution to the requirement. This is a perfectly normal and business-like approach. The situation at present is quite different. A telecommunications services supply monopoly is the norm in European countries.

Despite the current restrictions, however, the market for the supply of these services will rapidly become competitive. I am encouraged in this view by the recent publication of the CEC Green Paper on the development of a common market for telecommunications services and equipment, which clearly signals the need for such a competitive market-based approach, and by the CEPT/PTT plans for the development of what CEPT calls Managed Data Network Services (MDNS), and the proposals for 'one-stop shopping' for telecommunications services in Europe. Basically, the MDNS concept is that the PTTs will combine to provide customers with a complete data networking service — that is, they will build a private network for a customer from the component PTT services available. As a by-product, they will offer one-stop shopping where the customer can deal with a single national PTT for a trans-European network, rather than have to deal with many PTTs with all the complications that involves. CEPT has indicated that these MDNS services will be introduced in 1988, and it is expected that the details will be announced later this year.

These developments are encouraging. But they do not go far enough, and will not create the real competition in telecommunications in Europe that the customer really needs. It is important that customers have real freedom of choice to shop around for competitive prices and services — one-stop shopping is a great idea, provided one has a choice of shops. Real freedom of choice in telecommunications services will come, and sooner rather than later. It is essential, therefore, to look beyond the current PTT situation when planning longer-term strategy.

To better understand my concerns in this area, one has to take a closer look at the costs of basic data communications services in Europe. There are, at present,

two basic services provided throughout Europe. These are the X.25-based public packet-switched data network services (PSDN) which provide data rates at effectively 2,400 bits per second, and the analogue leased line telephone circuits for private networks, which support data rates of 9,600 bits per second. PSDN services are based on volume and duration charges, and are tariffed for low-volume traffic between terminals and hosts. At higher volumes, such as are typical between computers in the academic world, the current X.25 services become prohibitively expensive. Leased lines typically have an associated fixed annual rental charge, and no duration or volume pricing, because the customer is renting a point-to-point capacity, and the customer provides the switching of traffic between leased lines as required. Thus, for many networks, the leased lines approach is the most cost effective. However, in Europe, leased lines between countries are extremely expensive, many times more expensive per kilometre than within countries, and very much more expensive for comparable distances than in the United States. Thus, at present, building a European network using leased lines, while more cost effective in most cases than using X.25 services, is much more expensive in Europe than in the United States. Clearly, this situation cannot continue and the prices of leased lines will fall and be much more closely related to the costs. The CEC's Green Paper on telecommunications policy highlights these problems.

However, the situation is worse than it first appears. Some PTTs have proposed a volume- or time-related charge on leased line circuits, on top of the annual rental charges. The prospect of paying an annual rental for a fixed point-to-point capacity, and on top of that paying for the use of the circuit, does not gladden the average telecommunications manager's heart. When the volume charges are comparable to the PSDN charges, as is proposed, one might expect an unenthusiastic response from customers. Paying a volume-related price to a supplier whose costs for the service are fixed is hardly a move in the direction of competition. Unfortunately, while some PTTs have recognized the inherent absurdity of such proposals and have stated that they will not introduce such charges, I am still concerned that this view is not shared generally by all the PTTs in Europe. The problem with the high basic costs for international leased circuits also remains, and will have to be resolved — national boundaries within the European Community should no longer be used as an excuse for telecommunications tariff walls and high charges to customers.

The current situation with the national PTTs is a consequence of a multi-national rather than a European approach to the provision of telecommunications services. With the forging of closer links between countries clearly stated as Community policy, the telecommunications authorities are presented with a unique opportunity. A vast open market for products and services could be opened up, with the twin advantages of high volumes and increasing demand for specialized services. Yet instead of grasping the opportunity to create a European Telecommunications Authority (ideally, with several competing providers of services) the PTTs persist in a multi-national approach. Indeed, the recent CEC Green Paper fails to tackle this basic problem. The results for the European customers are obvious; charges are artificially high due to monopoly, there are tariff boundaries, over-regulation and duplication of administration, and customer choice is negligible. In addition, there are great differences in charges for PTT services in the various countries.

Although there is a lot going on in academic computer networking in Europe, most of the organizations involved take a traditional view of networking — that is, of users on large-scale computing systems connected to a wide-area network. In practice, most researchers are now using advanced work stations or personal computers for their work. In general, these work stations are connected together using some Local Area Network (LAN); in turn these LANs are connected together, and to the larger-scale computing facilities, by a Campus Network, using a variety of technologies and a variety of communications protocols. What is now required is that the researcher, through the single window of the local work station, via local, campus and national/European networks, should be able to gain access to required computing facilities and databases, and communicate with colleagues throughout the world.

My vision is not of one wide-area network, but of a vast network of networks interconnecting the researcher's local graphics work station to local, national and international computing resources, and enabling the researcher to draw upon powerful computational resources without explicit knowledge of the machines and communications involved. Furthermore, researchers should be able to access and interact with remote resources, as if they were local, and be able to take advantage of remote high-performance computing resources from the graphics display work station at their desks. The bandwidth requirements for such an advanced computing and communications environment are orders of magnitutude larger than currently available on the networks described above.

Such a high-speed network of networks, or Internet, is the model for academic networking in the United States, and is currently being implemented by the National Science Foundation's NSFnet programme. Such a model should be adopted immediately by the European academic community, and a concerted effort made to put it into practice over the next two or three years. This effort will require the cooperation of the current European networking organizations; the cooperation of the PTTs, so that high-performance (64 Kbit s^{-1} to 2 Mbit s^{-1}) circuits can be made available across Europe at prices comparable to those in the United States; the cooperation of the national networks; and, most importantly, the adoption of a European attitude.

To achieve such a truly European networking infrastructure requires, of course, a European source of funds. This necessity for European (as opposed to multi-national) funding is common to all such activities, and is the greatest political challenge to the Community. One is tempted to make a comparison with the Common Agricultural Policy (CAP) — although that may be politically unwise, it is valid when one remembers that the purpose of CAP was strategic, namely to make Europe self-sufficient in food. As such, CAP was enormously successful. Strategic objectives such as self-sufficiency in food, self-sufficiency and international competitiveness in microchip technology, competitive telecommunications supplier industries, competitive telecommunications services and an efficient networking infrastructure for academic research are as important for Europe as for the United States. They require European funding.

It can be argued that the multi-national approach to such objectives, through harmonization of national efforts, is the more appropriate way forward for Europe: indeed, in practice, such harmonization has been the thrust of the European Community political, social and economic activity for many years. This view can be and indeed is being challenged. The original concept of the six founding national signatories to the Treaty of Rome clearly envisaged the creation of a European Community, with its own structures and processes. This was to be the main weapon in the strategic battle to win back economic power to Europe from the United States and the rest of the world. Perhaps the recent agreement on the Single European Act demonstrates a return to the founding principles of a European Community, and the awakening of a need for a European perspective which transcends nationalism. Certainly in academic computer networking, and in telecommunications as a whole, a European, not a multinational, approach is required. □

Dennis Jennings, Director of the Computer Centre, University College, Dublin 4, Ireland, is the current President of EARN.

Reforming Soviet research

Soviet research is in good shape and even good spirits, but would be more effective if its managers paid more attention to the needs of its researchers and the framework in which they operate.

THIS brief survey of the state of science in the Soviet Union is informed by several prejudices, of which the more obvious are now listed.

First, science is international in the sense that even researchers of different nationalities and with different short-term goals (publishing a paper, making a bomb) have identical long-term interests (finding out). That is why it is has been a tragic loss for the West as well as for the Soviet Union that, for half a century, links between the Soviet enterprise in science and that elsewhere have been needlessly attenuated.

Second, the election of Mr Mikhail Gorbachev as General Secretary of the Communist Party of the USSR in March 1985 is the best thing to have happened in the Soviet Union, 70 years old next week, for at least 60 years; it may also the best thing to have happened to the world elsewhere since the end of the Second World War. Mr Gorbachev wants to reconstruct the Soviet Union and to make it prosperous, but to keep it socialist. Can anybody elsewhere pretend that the world would be a safer place if the Soviet Union were bankrupt and socialist?

Third, nobody can tell whether Mr Gorbachev's venture will succeed. The odds of history are against him; Russia changes only slowly. The fact that the intellectual community in the Soviet Union seems to be solidly behind him does not assure his success.The next few months, before the *ad hoc* party conference next June, will be critical.

Fourth, the Soviet government is (or should be) downcast that the attention given since 1917 to higher education, science and technology has left the Soviet Union with a civil industry which is inefficient and incapable of meeting social needs.

Science and technology have been cherished from the outset of the Soviet state precisely in the belief that only science and technology would create the resources required for the sustenance of a socialist state. Why has that calculation gone so wrong?

The achievements of Soviet science are a curious mixture of success and failure. Even before the revolution, Soviet science had made a distinctive mark (remember Mendeleev?). Since, there has been a long succession of distinguished Soviet practitioners and a long roster of practical achievements. Especially in mathematics and physics, but more recently in other fields as well, Soviet researchers have won great respect from their colleagues elsewhere. There are many ways in which the temper of Soviet science as practised at the best institutes and universities is directly in the tradition of old-fashioned European scholarship; undistracted by the scramble for research grants, unworried by doubts about their tenure of office and secure in the knowledge that their institutions will survive, able men and women devote themselves wholeheartedly to scholarship and the intellectual welfare of their students (also a part of scholarship).

Other Soviet researchers, equally distinguished and sometimes as well-known academically, have shown the Soviet government the recipes for making several startling machines that function successfully — nuclear weapons of various kinds, rockets that send satellites into orbit reliably and nuclear submarines apparently no less functional than those of other manufacturers (of which there are not many). Even those who wish these machines did not exist must agree that they play a crucial and legitimate part in the defence of the Soviet Union. Why cannot a technical enterprise so technically competent arrange for Moscow to have a reliable supply of toothpaste?

The failure to deliver the economic growth the Soviet state and now (with *glasnost* increasingly) the Soviet people look for is not the fault of Soviet science but of the economic environment. By declining to make science a scapegoat for obvious economic failure, all Soviet governments since the Second World War have implicitly acknowledged that. Mr Gorbachev's strength is that he has gone further than his predecessors in his appraisal of the magnitude of the Soviet Union's economic failure and in recognizing the dangers inherent in its continuation. He puts much of the blame on the shoulders of the 15 million bureaucrats, planners and decision-makers, intervening between the two halves of civil industry — production plants and their customers. What must worry him, as it does his supporters, is the difficulty of creating a link between the two halves that will give the former an incentive to produce what the latter want that will not be mistaken for ideological backsliding.

If Soviet science is not to blame, what can it do to help? Readers of the pages that follow will recognize that the Soviet research community abounds with talented people working imaginatively on important problems, just as in many other places. That impression is deliberately intended. But it should also be clear that the organization and management of Soviet science, unfamiliar to Westerners, also leave a great deal to be desired in Soviet terms. It is, for example, strange that a system attaching great importance to planning should allow the distribution of its talent among institutes and even geographically to be determined largely by the chance relationships of young men and women still at university with their teachers. There is indeed a danger that if the economic climate in which it operates were more efficient, Soviet science would not be able to meet the challenges that would face it. Moreover, the academy's programme of modest reform (see p. 781) is entirely insufficient to avert that risk.

Here is a brief list of some of the obvious defects of the present system, and some suggestions as to how, in Soviet terms, they could be remedied.

Mobility. Immobility within the Soviet research enterprise is stultifying. Outstanding people may be able to force their way to other institutes or even to persuade the authorities to let them create new institutes of their own, but most working researchers must reconcile themselves to staying for the remainder of their careers at the institutes to which they are first appointed. The University of Novosibirsk is even proud that half of all its physics graduates are still working in the town (most of the rest having been shuffled off into less exalted jobs). There is every reason to expect (and plenty of anecdotal evidence to confirm) that people thus placed are unlikely to be best placed, and are in danger of going to sleep, becoming malcontents or both. Academy administrators, saying there is "no problem", list the schemes by which people wishing to change jobs can, for example, look for others willing to swap apartments (a lottery that may succeed in 3 to 5 years or never). Within its own terms, it is mystifying that the academy does not follow other large employers of skilled labour in advertising all vacancies at all levels throughout its system and arranging that they are filled by merit from among all applicants, providing necessary assistance with relocation.

Deprivation. The days have gone when all practitioners of Soviet science were privileged. Now, only academicians and the similarly placed are privileged. Young people embarking on their careers are especially encumbered by the hardships of daily

Soviet life, but the establishment seems curiously indifferent to their hardship. Those who succeed against the odds are likely to be exceptionally able people (although nepotism may also do the trick, and often does, for friendships forged in adversity endure). But even those ground into the dust by adversity remain on the books. The official view is that, with last year's 25 per cent increase of university salaries (on a par with those of full-time researchers), "we are hoping it is our turn next". Should not a large employer of scientific labour anxious to get the best out of its workforce be more active in seeking the removal of impediments to its efficacy?

Injustice. Thoughtful enthusiasts for *perestroika* say that economic reform may be less urgent than social reform, especially that of the Soviet judicial system. A similar truth obtains within Soviet science. The most obvious cases of injustice are those in which Jewish people who apply for visas to emigrate to Israel usually (but not always) lose their jobs, in flat contradiction of the provisions of the Soviet constitution guaranteeing national equality (where a person's nationality is usually defined by the republic in which his parents were born, but for Jewish people is defined as 'Jew' and so recorded in his internal passport). A more common source of trouble is the arbitrary administration of the system for deciding who shall travel abroad, especially to the West. Although it seems that other government agencies and even organs of the party influence these decisions, they are administered (for academy staff) by the academy, which seems not to appreciate how arbitrariness can corrode morale. Few will deny the Soviet government's right to deny exit visas to some of those who apply for them, or the academy's right to decide when its employees can be spared to travel, but surely the government could insist that its nondiscriminatory laws are obeyed and the academy could arrange that its administration of its relations with the outside world is transparent and fair. The matter is the more urgent because of the shabby treatment of Sakharov in his exile.

Isolation. Much could and should be done to relieve the self-imposed isolation from which Soviet science suffers, and which explains why Soviet science is elsewhere less well respected than it should be. Simple stratagems would help, such as a courier service to the West to circumvent the delays that impede communications with Western journals. (Inward mail can take up to a month to reach its destination.) The bumbledom of the procedures for winning permission to publish abroad (increasingly a formality with *glasnost*) could surely be substituted by an institute director's fiat without compromising the Soviet Union's legitimate interest in commercial and military secrecy. More flexible policies on travel would plainly also help, although it is becoming clear that the cost of travel will be a handicap unless Mr Gorbachev can find a formula for currency reform that gives rubles real value. Other urgent problems of liaison are not even considered yet: what, for example, can be done to connect under-computerized Soviet science with the electronic data networks now well-established elsewhere?

Impoverishment. The performance of Soviet science is sadly hampered by equipment deficiencies on a scale too great to be cured by purchases from abroad, at least while rubles are what they are. Moreover, while the interest of individuals in research institutes is as variable as at present, it would be wasteful to arrange that everybody has the most modern tools for research. In the long run, this serious problem will be solved only when the ills of Soviet industry are cured. Meanwhile, there is an urgent need that the Soviet scientific enterprise should have more natural access than is provided by the often arbitrary allocations of hard currency to the international market in instruments and equipment. Plans that academy institutes with a present interest in the development of equipment should be free to make deals with companies overseas will unfortunately not produce results quickly. The capitalist solution of the problem would be for the Soviet Union to buy an instrument company elsewhere, using it as a means of selling its own designs internationally and using the proceeds for buying what it does not manufacture. Should not the academy be pushing for some solution of that kind?

Mismanagement. The sheer scale of the Soviet research enterprise means that its management must be a complicated task. The need that research institutes must be more self-reliant than in the West, ensuring that there is housing for their staffs for example, is a further complication. In the circumstances, it is inevitable that most attention should be spent on major projects, as in space research, plasma physics or high-energy physics.

By contrast, the management of more routine endeavours tends to be skimped. This, at least, is the best explanation why second-rate projects appear to survive indefinitely once begun, while there is a great deal of needless duplication within the system. (Most biology institutes, for example, support programmes of cell membrane research not easily distinguishable from each other.) It is asking too much of the academicians who function as heads of the academy's 17 divisions to deal with all these issues decisively. For the long run, there may be more promise in the scheme now being tried for supporting research in high-temperature superconductivity by asking interested laboratories to compete for funds. It will be interesting to see how this first national competition for research grants is adjudicated; as things are, the Soviet Union is the only major scientific enterprise in which such competitions do not exist.

Two other structural issues need attention, of which the chief is the conflict of interest arising within the academy's own operations. As things are, the academy is both an agent of research and a group of distinguished people which, naturally, has an important (and apparently welcome) collective influence within the Soviet government. Both functions are important, but not easily reconciled with each other. Saddled as it is with the research organization as it has evolved, the academy is not well placed to diagnose its faults and to remove them. It is also hamstrung by the plain truth that people chosen for their distinction in research are not necessarily able managers of the sprawling research enterprise. In the late 1960s, some academicians argued vigorously against the temptation to let the academy become the Soviet government's ministry of science, but events have gone against them. Now the academy is veering towards the election as academicians of people whose skills are managerial, which in the long run may undermine its influence and its usefulness. Either the academy or its masters should ask whether it would not be more effective if it shed its managerial responsibilities, concentrating instead on shaping the pattern of research, perhaps by the administration of a research grants scheme. Such a move would also diminish the often deadening influence of grand men in the Soviet scheme of things.

A better balance between the universities and institutes in the conduct of research is also urgently needed. The Soviet Union is unique (even compared with China) in its dependence on formal institutes rather than universities as the agents of research. Nobody would suggest that universities should have a monopoly of basic research, but there is now good reason why Soviet universities should have a better crack at the whip. The potential benefits for the quality of students' education should be evident, but it would help enormously to loosen the over-rigid framework of Soviet research if a measure of plurality were built into the system. Part of the present trouble is that the system is literally self-perpetuating; the numbers of young people qualifying as researchers is determined by the academic influence of those already working in the same field. In the circumstances, breaking new academic ground is bound to be difficult. It is remarkable that Soviet science has done so well. But would it not be even better placed if young men and women were not committed to some specialized field of science more or less from the beginnings of their careers? And is this not what the spirit of *glasnost* requires? □

Soviet academies

Coordination and management

THE Praesidium of the Academy of Sciences of the USSR is housed in the former mansion (now overrun by central Moscow) of a rich iron and steel family, with a handsome cast-iron staircase to advertise the family's wares. Three hundred metres away is a new white skyscraper (tall by Moscow's alluvial standards), with gilt moulding in a thick band around the top. The academy would have moved to the latter by now had the government not decreed that work on expensive buildings should be suspended.

Is the hiatus symbolic, perhaps of the uncertainty in knowing how *perestroika* will touch the most august of scientific institutions in the Soviet Union? Time, no doubt, will tell. Meanwhile, the academy has embarked on a modest programme of internally directed change.

That the Soviet academy is much more than a means by which distinguished scholars elect others of their kind to join them has its roots in the foundation of its predecessor, the St Petersburg Academy, by Peter the Great. His model was the Berlin academy of the eighteenth century. The academy has been an agent of research from the outset.

Now, with more than 250 research institutes and 200,000 research scientists on its payroll, the academy is not merely the agent of much of the Soviet Union's research effort but, because of its prestige as an institution and the personal influence of its members, has a dominating influence on the whole pattern of Soviet science and higher education. It is no wonder that the 250 academicians are widely respected and recognized as powerful people (but some academicians eschew the second attribute.) In social status, academicians compare with government ministers (of whom there are a great many in the Soviet Union.)

The growth of the academy's influence began with its move from Leningrad to Moscow in 1933, and rests to a large degree on the contribution of many academicians to meeting nationally important goals, the defence of the Soviet Union in particular.

More recently, academicians have played decisive parts in tackling national emergencies (vice-president Velikhov's role after Chernobyl), in shaping the pattern of industry (vice-president Ovchinnokov's influence on the development of biotechnology) and in developing a pacific space programme (as mounted by R.Z.Sagdeev's Space Research Institute).

Now, of course, there are other academies. Those for agricultural and medical sciences are national or "all-union" institutions modelled structurally on the academy of sciences. The fact that the academy was not able to halt Lysenko's influence on Soviet agriculture from his base at the academy of agricultural sciences may have seemed a black mark at the time, but his fall from power can only have helped the academy of sciences.

There are also academies of sciences for each of the fourteen republics of the union (apart from the Russian Republic), with a further 250 research institutes between them. For administrative purposes, the academy split off its Siberian activities into a separate division in 1958, and has created Far East and Urals divisions. Centres such as Leningrad, with a number of important research institutes, enjoy a degree of local autonomy similar to that of, say, the Far Eastern Branch.

Although these separate institutions all enjoy a degree of independence, their work is to a greater or lesser degree coordinated by what is generally called the "big academy", which also frequently takes the initiative in creating the committees coordinating other nationally important aspects of science.

This, for example, is the spirit in which the academy has this year taken the lead in safeguarding the Soviet government's interest in the development of ceramic superconductors. On the grounds that the first need is to understand the nature of the phenomenon, the academy has also broken new ground by inviting applications for research funds from all institutions considering themselves qualified to take a serious interest in the field, universities as well as public research institutes.

According to Dr Arnold Romanov, deputy chief scientific secretary (and a long-standing associate of Academician G.I. Marchuk, the newly-elected president), the new arrangement constitutes "an element of *perestroika*". It is too soon to tell, he says, whether this device for supporting research will be followed in other fields, but the State Committee on Science and Technology must first have a chance to give its opinion on the success of the scheme.

The management of the academy's interest in research is necessarily complicated. The academy is organized into 17 divisions, each of which is in the charge of an "academician-secretary" responsible for coordinating the work of relevant institutes and for organizing scientific committees reporting on special subjects. These, together with the vice-presidents of the academy and some 17 academicians without central executive responsibility, the president and the academician-secretary, constitute the praesidium of the academy, which meets once each week.

Individual academicians belong to one or other of the divisions, and in that role have been able to influence questions such as the appointment of people to fill important posts — directors of research institutes, for example.

The academy's relations with other organs of the Soviet government are also necessarily complicated. The academy is a creature of the Council of Ministers, essentially the executive branch of the government. The academy's work and future plans, like those of other academies, are grist for the mill of the State Committee on Science and Technology.

The academy is also the repository of a number of other important functions — the negotiation of agreements on the exchange of people with its opposite numbers elsewhere, for example. This is the spirit in which an academy delegation will be visiting Britain early in November, and in which a US delegation will arrive in Moscow later the same month. □

Acknowledgements

THIS survey would not have been possible without some of the changes Mr Mikhail Gorbachev has already engineered. The recipe is *glasnost*, meaning "openness", the buzz-word until it was overtaken in the spring by *perestroika*, meaning "reconstruction". (The Russian word for "reform" could have been used instead if it had not been discredited by its use in the early days of the Brezhnev era, now known as the "period of stagnation.")

The writer (John Maddox, editor of *Nature*), was invited to the Soviet Union as a guest of the Academy of Sciences of the USSR (referred to as "the academy" in what follows) and enabled to see what he had asked to see (with the exception of the Physical-Technical Institute in Moscow, said to be in a region of the city which is out of bounds to foreigners, and an appointment with vice-president of the academy, Y. Velikhov, which did not materialize). At no point did an academy scientist obfuscate the answer to a question, which is not to say that there were no disputed answers; some administrators were more reticent. Hotel and travel costs within the Soviet Union were met by the academy; *Nature* will find some way of reciprocating.

No spell of three weeks can condense into 24 pages a measured account of science in the Soviet Union — and this account is deficient in concentrating on the Russian republic (to the exclusion of the other fourteen) and on one of the several agencies responsible for research in the Soviet Union, the Academy of Sciences of the USSR (which is nevertheless the most important).

The writer is particularly grateful to his guides for most of three weeks. They are Professor Maxim Frank-Kaminetskii of the Moscow Institute of Molecular Genetics, a theoretical physicist turned molecular biologist, whose several applications to accept invitations to speak at conferences abroad have been rejected (most recently this summer), and who is Jewish, and Dr Anvar Shukurov, from the largely Moslem republic of Tadjikistan (on the border with Afghanistan), a graduate of Moscow University, and now a member of the staff of the Space Research Institute at Moscow, whose chief interest is in astrophysical magnetic dynamos. □

Academy reform

Modest changes in progress

THIS year will probably see more changes in the academy's way of doing business than there have been for several years. Among other things, steps are being taken to institute a procedure for electing the directors of research institutes and even the heads of institute laboratories.

The academy is falling into line with the Soviet government's prescription that there should be a formal retirement age for public servants. Earlier this year, for example, it agreed that members of its Praesidium should retire at 75 (roughly 14 of 40 exceed this age), but the rule will not be enforced until 1990.

In the same pursuit of youth, the election of new members of the academy due to be held in December will this time fill more places than will be vacated by the death of academicians, bringing the number of academicians up to 250 apart from those who exceed retirement age.

The electoral proposals have been stimulated by the Soviet government's plans for "democratization", and have been circulated to all institutes as a basis for discussion. The intention is that a final set of proposals should be ready for a meeting of the academy on 28 October, and that the new rules should be introduced on 1 November.

The draft reforms, as circulated for discussion, allow for the election of directors and laboratory chiefs every five years. Institute directors, at present nominated by the appropriate division of the academy, will in future be elected by that division from among as many candidates as there may be. Heads of laboratories will be elected by those working in the units concerned. But it is not yet clear whether the choice will rest with all the members of a laboratory's staff or merely with those who are scientifically qualified.

Similarly, it is proposed that some of the members of the scientific councils which formally govern the research institutes should be elected by the whole staff of the institute. One of the difficulties in this set of draft proposals is that scientific councils would become even larger than they are. It is not intended that the director of an institute should cease to be the chairman of his scientific council, while the draft proposals would return the right of appointment of laboratory chiefs to the academy if one of the candidates should be an academician.

Part of the controversy surrounding the draft proposals centres on the suggestion that institute directors should be elected by the appropriate division of the academy, not by the staff of the institute. Incumbent directors seem uniformly unhappy at the notion that their tenure of office is likely to be challenged as a matter of routine every five years; curiously, there is less discontent at the likelihood that they may not be in future personally responsible for the choice of those in charge of their laboratories.

The academy says it has also begun to give some of its institutes (such as the instrumentation institute at Leningrad) the freedom to engage in commercial arrangements with outside organizations. This is consonant with a general recommendation of the government that research institutes capable of making their way in the world by means of contracts with production ministries or individual factories should be required to become self-supporting. But president Academician G.I.Marchuk says that few of the academy's institutes are in that position, and that the government is committed to the continued support of basic research.

On the common complaint by researchers that Soviet research equipment does not compare with that of rival laboratories in the West, Arnold Romanov, deputy chief scientific secretary, acknowledges that there are problems but insists that the academy is not responsible for all of them. Thus, he says, in respect of computer equipment, the Soviet Union's time-lag behind the West arises because "the importance of this field was not appreciated" early enough by other agencies of the Soviet government.

Within academy institutes, researchers complain that the personal computers now coming into service tend to be unreliable; one theoretician with four such machines for use by his group says that he must reckon that at least one will be out of service at any time, partly because of the slow speed of service and partly because spare parts are only acquired with difficulty. But the most common comment on the computer problem, especially from theoreticians, is that the shortage has been a blessing in disguise, providing mathematicians with an incentive to improve on existing computational methods. □

New man in charge of academy

ACADEMICIAN G.I. Marchuk, elected president of the Soviet academy by secret ballot at the end of 1986, is a vigorous man who seems to enjoy his job. His presidency of the "big" academy seems almost a continuation of his long stint as president of the Siberian Division at Novosibirsk, although he spent two years as chairman of the State Committee on Science and Technology between the two posts.

His view of the academy's contribution to the changing climate is also cheerful — it should be the academy's job to identify the fields in which it already has a great deal of expertise and to push these along faster. But he also says that the changing climate should make industry more interested in the results of research, in which case there should be opportunities for the academy network to prosper.

Marchuk does not share the usual institute view that universities are a lesser kind of institution, saying that their staffs are as well qualified as those at the research institutes. The chief difficulty for the universities, he says, is that the production ministries which are the main users of university graduates in science and technology are unwilling to meet the full cost, with the result that universities in general are unable to establish themselves in research.

Marchuk is adamant that Soviet secondary education maintains higher standards than anywhere else "including the United States". Noting curriculum changes in the past decade, he says the objective now must be to strengthen the creative component of secondary education.

Marchuk is not greatly moved by the decline of interest in science and technology among young people, which he thinks "may not be altogether a bad thing". The causes, he believes, are simply that other occupations now offer higher salaries, while the hardships of the scientific life (and the interest of social issues) have inevitably changed the patterns of young people's choices.

Marchuk's arrival in his new office has not delighted all his fellow academicians. Some suspect the new administration's recruitment of what appears to be an alternative secretariat and are irked by the difficulty of reaching him informally by telephone. No doubt the greatest difficulty, at a time when change is almost a necessary precondition of self-respect, is to judge the pace and even the direction correctly. If anything, Marchuk may be over-cautious for the times. □

Government machine

Why the party calls the shots

MANY anti-communists in the West say that they would forgive the Soviet Union everything were it not a one-party state; only the Communist Party participates in elections. But the genteel sentiment that the system might be changed from within if the Communist Party were defeated at the ballot-box entirely mistakes the significance of the October Revolution, whose 70th anniversary is next week.

Remember the slogan "dictatorship of the proletariat"? How could it be assured that this hard-fought principle would continue to obtain if parties hostile to the principle were able to win power? Exactly similar fears of constitutional subversion have, from time to time, been used to justify the banning of communist parties in the United States.

Even so, people from the West must repeatedly remind themselves of this reason for the unfamiliarity of the Soviet system of government. The Communist Party (the "party" in what follows) is the inheritor of the revolution. Organized locally, regionally and centrally in a manner that exactly parallels the organization of local and national government, the party determines policy and monitors its implementation by the organs of government at all levels.

At the national level, the crucial organ of the party is the Central Committee. Its 300-odd members are elected every five years by the regionally representative party congress. Its staff of officials is the day-to-day monitor of the work of government and other public institutions (such as national newspapers and state-owned factories).

The executive committee of the Central Committee is the Politburo, whose 13 full members are inevitably powerful people. (One was retired last year.) It is natural that the general secretary of the party, now Mr Mikhail Gorbachev, while nominally first among equals on the Politburo, should be the dominant political figure in the Soviet Union: he controls the party machinery.

Local and regional party secretaries enjoy a similar kind of power within their own spheres of influence, and may exercise it arbitrarily or even corruptly. These are the grounds on which the senior party officials of the Asian republic of Khazakstan were replaced, earlier this year, by people nominated from Moscow.

Part of the objective of Mr Gorbachev's proposals in 1986 that there should be a choice of candidates in the election of local and regional party officials (by party members) is meant to weaken entrenched power at these levels in the hierarchy. His proposals for the 'democratization' of public enterprises, allowing workers in state-owned factories to elect their directors, would similarly ensure that local party organs do not have undue influence of appointments.

While the party and its organs are the repositories of political power, they have no legislative or administrative competence. At the national level that rests with the Supreme Soviet and the Council of Ministries (with a prime minister and too many ministers) respectively. Plainly the system would not work at all if certain key ministers (those responsible for foreign affairs and defence, in particular) were not also members of the Politburo.

The Supreme Soviet is a scrutineer of government legislation much as is the House of Lords in Britain. This machinery of government is mirrored in the arrangements at regional and local levels; ordinary people who are not party members have a vote for the membership of the local soviet, and are also now to be offered a choice of candidates.

Both the Academy of Sciences of the USSR (directly dependent on the Council of Ministers) and its analogues in the fifteen republics have traditionally been jealous of their distinctive procedures. Academicians and their officers are elected by the membership in secret ballot. Those seeking to be members or corresponding members of the academies may apply for vacancies in particular disciplinary divisions, but may also be nominated by the membership.

In principle, neither academicians nor academy officials, institute directors for example, need be members of the party, but practical considerations argue otherwise. First, party membership constitutes a kind of character reference. Second, a person with executive responsibility such as an institute director might well find himself at odds with his local party organization over policy were he not in a position to influence party opinion — though the other side of that coin is the risk of becoming a prisoner of unwelcome local party prejudices.

There is no doubt of the party's obtrusiveness even in administrative matters. In Leningrad, for example, the party takes the view that no institute scientist should make more than one extended visit overseas in any year. But party membership, increasingly a privilege, extends to only 10 per cent of the adult population.

There is nothing in what Mr Gorbachev has been saying in the past two-and-a-half years to suggest that he has doubts that the party should be supreme. *Perestroika* is about making government more efficient and the party more democratic. But it could emerge that the influence of the party on public (and even private) life could be modified without being accused of betraying the revolution. □

Press freedom

Glasnost may be only paper-thin

A FREE press is a remarkable thing. Most governments do not put up with the trouble of allowing one. The Soviet government just now is somewhere in between, recognizing that public discussion of important issues may be an indispensable curb on the entrenched power of local and regional party bosses, but still unsure how far *glasnost* should be allowed to go. Time will tell.

In the narrow sense, *glasnost* is an experiment being conducted jointly by the Soviet government and the Soviet press. The censors, who have kept their jobs, still see everything about to be published, but comment only if matters of 'national security' (whatever they may be) arise. The general press is still trying empirically to find out where to draw the line, which makes its exercise of its new freedom tentative and cautious.

MOSCOW NEWS

A difficult breakthrough to a world without medium- and shorter-range missiles

How to transcend the traditional holiday

UP THE GLASNOST STAIRCASE

Moscow News, *the Soviet tabloid making waves.*

In retrospect, *glasnost* may be a consequence of the accident at the Chernobyl nuclear power station in April last year; 280 million citizens (not to mention neighbours) had to be reassured. In August that year, Mr Igor Yakovlev was appointed editor of *Moscow News*, a weekly previously put out through Intourist hotels to help tourists decide how to fill their time.

Moscow News is now a very different tabloid, bursting at the seams with political matters. Sales of the Russian-language edition (on Tuesdays) bump up against the ceiling of 250,000 copies supposedly set by the capacity of *Moscow Pravda*'s presses. Exact translations appear in half-a-dozen other languages on Thursdays. Its partner as the spearhead of *glasnost* is the

Censorship makes headlines in *Ogonyok*

In at least one crucial respect, *glasnost* has brought a change of policy on the censorship of journals such as *Nature*. The old practice of photocopying the journal ceased when the Soviet Union signed the International Copyright Convention two years ago. Then selected news articles from individual issues would be omitted from the Soviet version.

Until February this year, some copies of *Nature* received from Britain would be censored before reaching library shelves, in at least some institutes by removal of entire pages carrying offending articles. Brief inspection suggests that the censorship was not especially subtle; articles critical of the US Strategic Defense Initiative seem on some occasions to have been removed, for example. Much the same seems to have happened with the comparable journal *Science*, which is reproduced under licence in the Soviet Union.

Why and when the practice ceased is not especially clear. At some institutes, all copies of *Nature* appearing since February have been left intact, but photocopies of contents pages distributed by the Institute of Scientific Information of Philadelphia carried blacked-out entries as recently as April. Academician V. Goldanskii has brought the issue to the attention of a wider public in a letter printed in the widely read weekly magazine *Ogonyok* ("Little Flame"). Is he pushing at an open door? □

weekly general-interest magazine *Ogonyok*, which (among other things) serializes previously banned literature, but which published a protest by Academician V. Goldanskii at the Soviet practice (now apparently defunct) of censoring the content of journals such as *Nature*.

It is significant that the deputy editor of *Moscow News*, V.A. Buzchkin, should have acknowledged earlier this month that even his courageous editor will telephone some official at the Central Committee before running controversial matter. But the precaution is not always a safeguard; when *Moscow News* last month published an obituary of a Soviet emigré writer, there seems to have been a row between the editor, Mr Yakovlev, and Mr Igor Ligachev, number two in the Politburo after Mr Gorbachev himself.

Questions of the administration of science are now being discussed in the general press. Last month, for example, *Literaturnaya Gazeta* published a protest at the academy's failure to decree that even institute directors should be freely elected by secret ballot (see *Nature* **329**, 193; 1987). At about the same time, *Moscow News* on 5 October, referring to an earlier article asking why the Soviet Union derives so little benefit from its huge investment in research, concluded that matters would be improved if factory managers used market principles to commission research and development directly from research institutes of their choice, instead of having to go through the central bureaucracy. The previous week, the newspaper had given a graphic account of the circumstances (see p. 794) leading to the closure of the economics institute at Baku by the Minister of Higher Education. (The institute was shot through with nepotism in both the appointment of teachers and the award of degrees and diplomas.)

Despite this heady stuff, the academy's own science publications such as *Priroda* (which means "Nature"), *Science & Life* (with sales of more than 2 million copies a month) and *Chemistry & Life* (more than 200,000 copies) have so far not been much changed. Moreover, while the general press is still tentatively exploring its new hunting ground (and the limits of its freedom), its coverage of professional issues is at best fitful.

This is *glasnost* in the narrow sense. Beyond journalism, the general lifting of restraint in private conversation may in the long run be more important. People say that colleagues still harbour some of their old caution in dealings with each other. It is understandable that many may find it easier to talk openly to foreigners, but this foreigner encountered only half-a-dozen people in three weeks who were evidently telling sugar-coated tales. One, explaining why *glasnost* had not affected his behaviour, said "I've always said what's on my mind, but that doesn't mean that I tell everything that's there".

Even so, the optimists believe that *glasnost* has now gone so far as to be irreversible. They rely on the principle of the Tree of Knowledge. Pessimists, many of whom are also optimists, note that the censors are still at their posts, ready to pick up their telephones at short notice.

It is relevant that many of the administrative controls on the freedom of Soviet citizens also remain in being. Internal passports must be shown repeatedly, Soviet citizens are arbitrarily prevented from entering the hotels set aside for foreigners. Leaving the country even briefly remains a major undertaking. Incoming mail is still censored. External communications are difficult and slow. Even in foreigners' hotels, the only English-language periodicals available are the British Communist Party newspaper *Morning Star* and its equivalents from other European countries.

Glasnost, meanwhile, has lifted the spirits of scientists in the Soviet Union. Many are childlike in their enthusiasm for a more open press. Few have yet learned that *glasnost* is also an invitation for them to speak out. □

Perestroika

Economists come in from the cold

The question "What are you doing for *perestroika*?" may be a reporter's cop-out, but not the less useful on that account. Most researchers in the natural sciences reply by saying that they continue to work hard in their laboratories, believing that to be the best contribution they can make. Occasionally they go on to say what they hope *perestroika* will do for them (make it easier to buy a cooperative apartment, for example).

Although the intellectual community seems solidly supportive of the prospect of change, the new climate seems not yet to have reached the point at which more than a handful of researchers believe they can by persuasion improve the conditions under which they work.

But that is not true of economists. The changed climate has brought them out of the shadows in which they have been hidden since Lenin's New Economic Programme of the early 1920s. That was the period of emergency when entrepreneurs were allowed to make profits, sometimes with the help of Western capital.

Western economists are forever saying that their forecasts are unreliable because their assumptions are necessarily assumptions about human behaviour, in its nature unpredictable. They may be chastened to know that there are many in the Soviet Union who believe that the Soviet economy responds only slowly to stimulus because of a unique and admonitory influence on the behaviour of would-be risk-takers in the Soviet Union: when Stalin unwound Lenin's New Economic Programme, he arranged for many of those who had prospered to be shot.

But Soviet economists seem now prepared to take a chance. The best-known of the adventurers is Academician A.G. Aganbegyan (see p. 800), who seems well on the way to having a mathematical model of *perestroika*. Dr Andrei Belykh, of the University of Leningrad, while sharing the general enthusiasm for change, stresses that excitement about the prospect of change is not new in the Soviet Union.

Did not both the Khruschev and Brezhnev eras begin with promises of economic growth? He also notes that the first of Brezhnev's five-year plans (1965–70) was the only plan since the revolution to show an acceleration of the annual rate of economic growth, which otherwise has usually worked out at between 3.1 and 3.3 per cent. Surprisingly, he also refers to a paper by Stalin in the early 1950s ("The problems of Socialism in the USSR"), which tacitly acknowledged that prices must be some-

how linked to costs. Brezhnev's reforms of the 1960s failed, he says, because they were exclusively economic.

According to Belykh, the immediate economic goal of *perestroika*, that of moderating consumer subsidies and allowing salaries to rise in compensation, has been on the planners' agenda for well over a decade. Of the logic, he says, there is no doubt, but the speed of change remains in doubt.

Aganbegyan has been saying that an economic relationship between costs and prices should be in place by 1991, but Belykh believes that may not be feasible. The immediate difficulty is that Soviet economic statistics are inadequate for an accurate calculation of the trade-off. "It would be unwise to say that our statistics are in an exceptionally good condition."

The extent to which the lack of statistics may cramp Mr Gorbachev's economic reforms may not yet be fully appreciated in the West or even in the Kremlin.

Although the Soviet government has access to a wealth of data about the production of goods of different kinds, there is no means of telling how people's patterns of consumption would change if prices changed. In a capitalist economy, this would not matter; the market would bring equilibrium. The danger, in a planned economy is that ignorance may breed inaction.

Belykh is also open-minded about the likely consequences of the process of democratization, the election of factory and other managers; "the proof of the pudding is in the eating". If the new system does not work, "we'll have to find another", he says. Nobody can tell what the future holds.

But there is one certainty, in Belykh's opinion: the Soviet Union will never countenance unemployment. But it may at some stage be prepared not to pay high wages to those who do not wish to work, allowing only the productive to share in rising prosperity. Some would say that is a prescription for a welfare state British-style (*circa* 1968).

Meanwhile, it seems that academic economists are being consulted again by government at all levels. Belykh believes, moreover, that he and his colleagues have a great deal to contribute. The trouble, it seems, is that the pace of change dictated by political considerations may be too quick. □

Superconductivity catches *Pravda's* eye

On "Scientists' Day" in April this year the Soviet Union's contribution to the research topic of the moment was reported in glowing terms. Since then, the academy has broken new ground in its mechanism of research grants from its own institutes and other research institutions. Scrutineers say that one award will certainly go to the institute at Tomsk (Siberia) that has succeedeed in making twin films of ceramic materials.

ПРАВДА

Сегодня—День советской науки

ПРОРЫВ В СВЕРХПРОВОДИМОСТЬ

А. ПОКРОВСКИЙ.

На снимках: в лаборатории сверхпроводимости — заведующий лабораторией А. Голованцев и научные сотрудники К. Ищук, Е. Пичкин и О. Мазанко проводят эксперимент; в одной из лабораторий института. Фото К. Кардонова.

Fiscal facts

Money means more to capitalists

FOREIGNERS, the Soviet term for people from elsewhere, should beware of the Soviet concept of money and of the frequently arbitrary values attached to it.

At the official counters, £1 will buy 1.05 rubles; $1 will buy 0.65 rubles or 65 kopeks. But the equivalence is entirely artificial. One can travel 30 km on the Moscow metro for 15 kopeks, buy a decent breakfast at a hotel buffet for about a ruble or a good seat at the ballet for less than 3 rubles. Domestic air travel is cheap (100 rubles will take a person from one end of the Soviet Union to the other) but tickets hard to purchase; ordinary travellers are allocated seats only when official travellers and tourists (who pay more) have been accommodated.

But some things cost more than elsewhere. International telephone calls to a nearby country such as Britain seem to cost about six times the external rate, and Russian vodka is about the same price in Moscow foreign currency shops as in highly taxed Britain.

Part of the explanation is the Soviet Union's way of socialism. The necessities of life are sold to Soviet citizens at subsidized prices (to which foreigners are properly denied access by an artificially high exchange rate), but luxuries are charged for through the nose, so only foreigners can afford them.

That the exchange rates are artificial is tacitly acknowledged by the procedure under which Soviet citizens with foreign earnings, say from authorship, are allowed either to keep them in a special bank account against future foreign journeys or to convert them, at a ruble rate of 4.6:1, into 'checks' used for the purchase of luxury goods. The checks are convertible (but illegally) into paper rubles at the ratio of 1.7:1, suggesting a 'market' conversion rate of 8 rubles to £1 or thereabouts.

The other side of this coin is that employment is also rewarded cheaply. Wages and salaries are uniformly low. Moscow bus-drivers now earn 350 rubles a month, more than twice as much as 'junior scientific workers' (people with the Soviet equivalent of a PhD at the beginning of their careers). Senior scientists with a BSc may earn 500 rubles a month. Members of the academy earn an extra 500 rubles on that account and a further 150 rubles a month if they become members of the academy's praesidium.

If the objective were simply to match low prices with low salaries, nobody would complain. What Soviet citizens do complain about is that there is little in the shops they want to buy, while services (such as restaurant meals) are almost equally 'unavailable'. The result is that many people have money in their pockets (or in savings accounts earning 2 per cent interest), but are perpetually frustrated by the lack of material goods or even of good food and drink. The free markets for high-quality vegetables and better-quality clothing in the large cities seem not to have dented this mismatch between the over-supply of money and the frustrated demand for goods.

Not that young people share this view. Their chief discontent is housing, or the general lack of it. Nominally, housing costs are low. A rented municipally owned apartment will cost a few rubles a month, the purchase of a share in a cooperative apartment block somewhere in excess of 1,200 rubles and a privately owned house (there are such things) ten times (or much more) as much.

The snags are that waiting lists for municipal apartments are several years long, at least in the major cities. The allocation systems are skewed in favour of party members and other dignitaries and even the apparently modest cost of cooperative apartments is beyond young people with only modest savings or with-

out indulgent parents. Motor cars seem further beyond reach.

There is no mechanism for providing private persons with consumer credit. As a consequence, the Soviet government cannot control the level of personal consumption by controlling interest rates, a device considered essential by Western governments. So, in Western terms, the Soviet Union is in perpetual deflation; the money that people save under their mattresses against an uncertain future is money taken out of the economy.

This state of affairs is distinctive to the Soviet way of socialism. To make life's essentials (food and housing) and good things (books and ballets) cheap may ensure that they are in principle available to all (but then there may also have to be an allocation system, which may be operated unjustly). The economic costs of these subsidies are sometimes (as with the Leningrad metro) never calculated. Sometimes, subsidies are not rationally related to each other, so that people empty the bookshops of books they will never read or (in the country) feed 15-kopek-a-kilo bread to livestock.

That this system has required the Soviet Union to insulate itself financially from the outside world, making the ruble unconvertible against Western currencies, always an administrative nuisance, has become an economic burden.

Mr Gorbachev's still undefined programme of reform will be incomplete without some kind of monetary reform. He told a group of workers at Murmansk just three weeks ago that prices would have to rise. The planners are apparently looking for ways in which salaries will also rise in compensation, but it is difficult to see how that can be done equitably. Mr Gorbachev will probably have to invent a domestic banking system.

One of the obstacles seems to be lack of economically significant data. Not that the bureaucracy is short of numbers; the problem is to relate them to economic costs. Another is the Soviet attitude towards money, which long predates the revolution: as one distinguished Soviet scientist said, tongue in cheek, "we are supposed to devote ourselves to science, and to believe that brings its own rewards".

This other-worldliness seems to have led to the view that money is merely a way of simplifying the old barter system, and that it is an inappropriate measure of economic costs and benefits. The directors of quite large enterprises are often surprisingly vague about the details of their budgets; is it any wonder that the all-powerful party, working on the principle that the government is merely its servant, will often will the end of great objectives, leaving the bureaucrats to find the financial means? Mr Gorbachev will have to create a tough-talking treasury department at some stage. □

Great men and barons

Differences of style in science

EVERY Wednesday morning during the academic year, Academician V.L. Ginzburg, just turned seventy, opens his public seminar at the Lebedev Institute in Moscow. Anybody is free to attend. By everybody's account, it is an occasion of a kind unknown in the West. Ginzburg says what he has found interesting in his reading, other people do the same and then the discussion centres on some subject on which he has something special to say. That may be any theoretical topic in physics, but given Ginzburg's special interest in superconductivity, the discussion these days is usually about the mechanism of superconductivity in the new ceramic materials — and the audience frequently overfills the institute's largest auditorium. Each seminar finishes promptly at one o'clock.

This occasion says something about the distinctive style of Soviet science. The great men diligently practise the socratic method, willingly walking the tightrope of unscripted public discussion with their senior colleagues for the benefit of students, living by their wits in public. Their model is the late Lev Landau, tragically incapacitated in a motor crash, who died in 1968, leaving behind what is regarded as a masterpiece of a textbook.

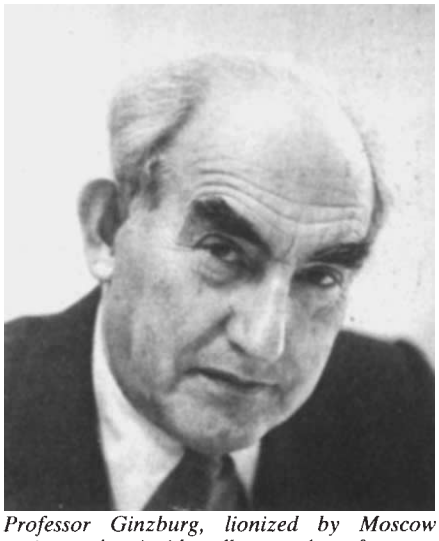
There may be a handful of people in the Soviet Union who can sustain such an ordeal. Another, for example, is Academician Ya B. Zeldovich, with Sakharov responsible for the Soviet hydrogen bomb, who conducts a weekly three-hour seminar on astrophysical topics. His students are told they must attend even when they know too little to follow the proceedings. Sooner or later, they will be asked to discuss a paper Zeldovich considers important at the following week's seminar.

It is no accident that these best-known seminars are by theoretical physicists, for their field is both that at the pinnacle of Soviet respect for science and the starting-point of the education of a large proportion of all Soviet scientists. It is the basis of the reductionist style of Soviet science: is not 'making science' for explaining the real world in terms of elementary concepts, after all?

So the ideal piece of work is a fresh hypothesis, a theoretical exploration of its consequences and a conclusion that coincides with observation or, sometimes, merely common sense. Even those working in much more empirical fields, instrument designers for example, worry away at the details of the principles on which their instruments are designed. There is no reason to believe that easier access to first-class equipment would change this way of working. There are obvious benefits in a system ensuring that many experimentalists have a strong background in theory. The benefits for reporters are considerable: people's explanations of their work begin at the beginning.

One side-effect may be the reason why many Soviet papers appear excessively theoretical to Western readers (and journals). The same tradition, largely a successful one, also colours the way in which students are taught. Nobody should be surprised that the introductory mechanics course at Novosibirsk State University begins with special relativity, for example. Another incidental consequence is that Soviet molecular biology is well populated by migrants from the physical sciences (chemistry included).

Two other features of the system further help to make the Soviet style distinctive, of which the most significant must be the intense personalization of Soviet science; people's relationships with others



Professor Ginzburg, lionized by Moscow society and, coincidentally, coauthor of a paper in this week's Nature. *('Cosmic rays and gamma radiation from the shell of supernova 1987A', page 807.)*

— teachers, students, co-workers, not to mention institute directors and academicians — appear more obtrusive than in the West. In part, this seems to be because of the way that, in such an army of researchers, strength of personality seems to be a necessary (although not sufficient) condition for success.

The immobility of the system, which binds groups of people together for years on end, and the way in which the educational system attaches young people to particular scientists from a relatively early age, reinforce the personalization of research. After three years of university education, good students enter an academy institute (less good students enter less

highly regarded institutes or even teacher training institutes) and, after a further year, may be attached to a particular senior person. The better they are, the more likely it is that their relationships will persist — perhaps for a decade or more.

One gossipy result is that people surprisingly often know that some rising star was a "pupil of so-and-so". Another is that great men tend to accumulate extensive groups of long-term collaborators, possibly straddling several institutes, which then become 'schools'. The prizes, for those at the hubs of such networks, are the opportunity to be productive, the intellectual reputation that follows and even the chance to set up a separate institute.

People like Ginzburg and Zeldovich, knowing the respect in which they are held, can afford not to allow external responsibilities to distract them from research. (Ginzburg, for example, hopes to take advantage of the new rules in the academy to cease being head of the theoretical physics laboratory at the Lebedev Institute early in the new year.) Even so, for people of distinction, teaching responsibilities can be onerous.

The darker side of this tendency for important men (and a few women) to occupy dominant places in the pattern of Soviet science is the temptation for people who are less than sure of themselves intellectually to simulate the pattern characteristic of success by the accumulation of public offices — memberships, or chairmanships, of academy and government committees, posts as editors of journals and the like. This pattern is well described by a passage from a complaint in *Moscow News* from one Sandro Belotsky, a physician, at the growing importance of the "science manager" in the Soviet Union.

> It is no secret that, in this country, promotion is dependent on the individual's social activity. There's nothing fundamentally bad about it, but the system does have one flaw: an active worker inevitably accumulates duties, which inevitably become posts. As a result, the director of an institute may at the same time be the head of a [university] department. . ., an academician, a member of the learned council of a ministry, the editor-in-chief of a scientific journal, chief consultant to a ministry, the chairman of a board of directors and member of the higher competence test commission [awarding higher degrees]. . .

Power-brokers of this kind do seem to play an influential part in the affairs of Soviet science. They are in no sense incompetent people but, rather, are often good scientists. But they can frequently be over-stretched and, in that condition, locked into symbiotic relationships with dependent henchmen in the discharge of their responsibilities. In the pursuit of loyalty and the discouragement of disloyalty, power may be exercised arbitrarily. Often, the barons are powerful enough to fight through the bureaucracy to accomplish tasks on behalf of science that would be beyond ordinary mortals. Lysenko was such a person.

This phenomenon is not exclusive to the Soviet Union, of course; the late Detlev W. Bronk, the first president of the Rockefeller University in New York, was such a person. But power-broking on this scale is less common in the West than in the Soviet Union, perhaps because of the hierarchical character of the Soviet system, perhaps from tradition or perhaps because, until recently (but see p. 782), advancing age has not been a bar to public office. As everywhere, the phenomenon is particularly distressing when it leads great men to scientific folly, or to add their names to the better papers of their junior colleagues. □

Moscow's environs

How to beat the bureaucrats

PROFESSOR Anatoly Logunov, rector of Moscow University, a vice-president of the academy and the chairman of the council that calls the shots at the high-energy physics research establishment at Serpukhov, asks a simple question: "What would you do with the more than a hundred research institutes in the Moscow region, each with between 100 and 150 people, with few PhD scientists and with an investment of a few hundred rubles per scientist?" Why not shut them down? Logunov shrugs his shoulders, mentions the difficulties and refers to the "bacillus of inability" that afflicts the Soviet Union.

Logunov was referring to research institutes of agencies other than the academy. His question was a rhetorical reply to the question how the academy itself would improve the quality of its own substandard research institutes. An impatient man, he made no secret of his own wish to see some action — not the "illusion of action" that had become fashionable during the Brezhnev "period of stagnation". Given the chance, he too would shut down all these institutes.

In many spheres of Soviet life, Logunov believes, the "bureaucrats have become more powerful than the directors" or policy-makers. He offers the academy network, and his own efforts to persuade the Central Committee to sanction Serpukhov's new machine, as an illustration of how determination can cut through bureaucratic fog.

Earlier, Logunov (now 60) had sounded vaguely defensive on the issue whether the academy should enforce its new strict rules on the retirement from public office of ageing academicians, explaining that "that would soon affect me" and that people were naturally tempted to believe that, the older they became, the wiser their decisions must be.

He is surprisingly outspoken about the universities and military service, even though science students at his own university are exempt from the Army. He would like to see the system of conscription changed for other institutions as well.

Logunov is a veteran of the famous visit to Britain by Kurchatov in April 1956, at which the Soviet delegation gave the news of their hopes for thermonuclear fusion. He is now temperately resentful of the way in which journals in the West, and especially those who act as referees for them, deal with articles submitted from the Soviet Union. As power-brokers go, Logunov is among the most congenial. Given the chance, he would make Soviet science much more efficient. But he has recently blotted his copybook with his less worldy colleagues with what is held to be an eccentric theory of gravitation offered as an alternative to Einstein's general theory, from which black holes among other things are abolished.

There has been much controversy since an account of the theory appeared in *Science & Life*; a more formal version appears in the proceedings of the VIIth international symposium on high-energy spin physics, at Protvino a year ago, and now published (Serpukhov, 1987). Some fear Logunov might become the physicists' Lysenko, but that seems improbable. □

The biggest university building in the world.

Sakharov

No more fences about Sakharov

ANDREI Sakharov is a saint, his friends say, but he is a saint in crumpled trousers, a well-used lumberjack shirt and a kind of cycling jacket made of wine-red velour which he wears open at the neck. It is easy to forget that he was the one who, with Ya B. Zeldovich, sketched out the principles on which Soviet hydrogen bombs were developed when he was a mere 32, and that he made his first protest about the international arms race as far back as 1961.

Since Sakharov and his wife Elena Bonner returned from his seven-year exile in Gorkii in December 1986, they have lived in an apartment near the Kurskii Railway terminal occupied by Mrs Bonner's mother, who spent part of her daughter's exile in the United States, but who has now returned.

Sakharov and his wife are late owls. He works at the Lebedev Institute, all day, relaxes over dinner, and then starts work again at nine o'clock or so. IBM may have influenced his working pattern by equipping him with a late-model personal computer, complete with one of the few personally owned printers in the Soviet Union.

Sakharov and his wife, Yelena Bonner, soon after hearing that their seven-year internal exile was to end.

Sakharov appears entirely without guile, willing to tackle even the most direct questions directly. It is almost as if he feels he has an obligation to answer any question put to him. He likes to develop his opinion, which is natural enough in one kept to himself for seven years. His friends offer a running commentary on what he says and his wife makes splendid coffee in the tiny kitchen in which people also eat.

Sakharov is, to say the least of it, disappointed that his fellow academicians did so little to help him during his exile. It would not be an exaggeration to say that he is bitter about the whole business. It is particularly hurtful to him that four members of the academy should have denounced him vituperatively in a letter published in the newspapers. He can understand why Academician G.K. Skryabin, the academy secretary, signed the letter, but why should A.M. Prokhorov, otherwise a decent fellow, have added his name?

Another source of discontent is that there is no sign that the authorities plan to return the medals he was awarded (with Zeldovich) for his work on the development of nuclear weapons, of which he was stripped at the outset of his exile. Zeldovich, sporting an exactly similar set of medals on the first day of the sputnik celebrations earlier this month ("We were told to wear them"), explained their absence on the second day by saying "How can I wear them in the presence of Sakharov, who is equally entitled to them?".

Sakharov is also mildy resentful (but not particularly surprised) that the video film of him in apparently good health, which was widely distributed in the West in the months before his exile ended, had been made by the use of the "candid camera" techniques of hidden cameras.

On the state of Soviet science, Sakharov says it is no surprise that theoretical research is stronger than experimental work; the shortage of equipment is a sufficient explanation. But he agrees with the opinion of another guest that Soviet physicists often drop projects when it would be more profitable to follow through, as with the discovery of superfluidity in liquid helium-III in the past decade.

He is also worried by the wayward tendency among Soviet scientists to follow eccentric leads and also to tolerate eccentric leaders, of whom Lysenko ("far from being a simple fool") is the best known.

Sakharov recalls with a blend of wonder and amusement the troubles in the early 1950s over the interpretation of quantum mechanics; Bohr and all his works at Copenhagen were then roundly condemned as idealists (then a term of abuse) by people such as Terletsky and Ivanenko. In the end, it took the intervention of Kapitza to quash the dissident movement. Now the Soviet scientific community as a whole is probably too jealous of its reputation elsewhere to tolerate waywardness on a scale that matters.

The need now, Sakharov believes, is for more durable contacts between the Soviet Union and the outside world — "uncontrolled migration", he calls it. But creative contact with the outside world cannot be maintained by "business trips" and, it goes without saying, "those who are invited" from elsewhere "should be those who are sent". "Nobody can tell who will bring the new ideas".

On the future of arms control, Sakharov is surprisingly cautious. He supports Mr Mikhail Gorbachev's initiatives, but emphasizes that the motives are domestic, mostly economic. "He has to release resources." He welcomes the prospect of an agreement on missiles of intermediate range, but is anxious that this should not provoke a loosening of US influence or a decline of US interest in Western Europe. (Mrs Bonner quickly dismisses as a "Western idea" the suggestion that the other side of that coin might be lessened Soviet zeal in controlling what happens in Eastern Europe.)

Sakharov is looking for other kinds of international agreements, not yet much discussed. On the grounds that the world will need nuclear power for the next century, but because of the consequences of the accident at Chernobyl, he would like to see an international agreement to put all reactors underground. He would also like to see a comprehensive test-ban, to which there is now no obstacle.

Sakharov's release? Mikhail Gorbachev's call with the good news came on a newly installed telephone.

For the rest, in this small kitchen to which the whole world seems to be beating a path, there is some casual conversation about general relativity, speculation about what may lie beneath, or beyond, the much publicized row between a member of the Politburo and the editor of *Moscow News* (see p. 783), one copy of which is passed round the table.

Only when trying to persuade a taxi to go in the right direction did our fellow guest explain that he had not visited the Soviet Union (where he had been born) since Sakharov was locked up. "He really is a saint", he said.

Sakharov may be glad to be away from Gorkii, but is not now cheerfully looking forward to an unclouded future. There may be promise ahead, but there are also great dangers, especially for saints. □

High-energy physics

Friends in high places

HIGH-energy physics is riding high in the Soviet Union, since the Central Committee announced on 17 July that the government would indeed provide the funds (in round numbers, 1,000 million rubles) for building a superconducting 3,000 GeV (3 TeV) proton accelerator. Eventually, the machine may be capable of engineering collisions between protons each of which carries 3 TeV of energy.

The result is that the management of the institute has been galvanized by the transition from exploratory work on the new machine to detailed preparation for the long haul ahead. (The first phase of the new machine will not be commissioned until 1993.)

Several kilometres of the 21-km long tunnel, at 5-m diameter big enough to take a single-track metro railway, have already been excavated. But the most urgent need seems to be that of housing the extra workforce being recruited, with the result that much of the site, splendidly landscaped with the mixture of birch and pine that makes people weep for Mother Russia, has been turned into a quagmire.

One remarkable feature of this development is that it appears to have been decided without the arguments there would have been in Western countries about the relative benefits of investment in high-energy physics and in other fields of science.

The Central Committeee, although not noted for its knowledge of high-energy physics, apparently shares the institute's view that all fundamental science is important, that high-energy physics is especially fundamental and that big machines are both a spur to technological development and a means of attracting international collaborations. Last July's decision seems to be yet another illustration of how decisions on major projects in the Soviet Union are made at a higher level than that at which funds must be provided.

Understandably in the circumstances, there is some amusement at Serpukhov, 90 km south of Moscow, that a decision is still awaited in the United States about a site for the Superconducting Super Collider, and that Britain appears to have cold feet about continued membership in CERN, the European high-energy physics laboratory at Geneva. No doubt it helps that the Institute of High Energy Physics is a dependant of the State Committee for the Utilization of Atomic Energy.

The plan is ambitious, but based deliberately on cautious engineering. The circular tunnel may eventually accommodate three accelerators. For the time being, funds have been approved for two of these, a 400 GeV accelerator (with conventional magnets, now called 'warm') and a 3 TeV machine with superconducting magnets. The institute will go back to the government in time for a decision five years from now on the installation of a second superconducting accelerator.

Interestingly, not everybody is persuaded of the wisdom of this strategy. At the end of last month, Professor Nicolai Dikansky, from the nuclear physics institute at Novosibirsk (see p. 790), broke a journey to California to visit Serpukhov so as to urge the planners that they could do without the second superconducting ring by relying on the techniques of electron cooling to accelerate antiprotons as well as protons in the first ring.

Professor Victor A. Yarba, vice-director of the institute, says that the argument has been settled, and that Serpukhov has decided to build a machine it knows will work. But there is every likelihood that the argument will continue. Five years is a long time, even in high-energy physics.

Whatever the outcome, the existing 70-GeV proton accelerator will be upgraded, by replacing the vacuum chamber, reducing the operating vacuum and replacing the injector, to secure a fivefold increase of the number of protons in each 8-second cycle. (Each bunch will contain 6×10^{14} protons.) Bunches from this machine will then be fed into the warm ring in the large tunnel and accelerated between 400 and 600 GeV, transferred to the superconducting ring and accelerated to 3 TeV.

As with all synchrotrons, each fully loaded ring will be able to function as a storage ring, keeping bunches intact for use over a period of some hours. Collisions between particles will be arranged by extracting appropriate bunches and bending their paths by means of electromagnets over a distance of 250 m or so, but the complications of that arrangement will no doubt be more clearly in the minds of the machine-builders in 1993, when the new machine is due to be commissioned.

More immediately, people are much more preoccupied with the manufacture of the superconducting magnets, each of them 6 m long. Some 2,000 will be needed if both superconducting rings are built. The ring design supposes a field strength of 5 tesla (50,000 gauss) in the vacuum chamber from conventional (not ceramic) superconductors. The windings are embedded in resin and baked; the designers hope that they will then retain their positions to within 25 μm, essential if departures from uniformity of the magnetic field are to be kept to within one part in 10,000. These assemblies are then embedded in stainless steel and the whole is prestressed in a hydraulic press built in the institute's workshops.

Industry, says Yarba, would have have taken three years to build and deliver such a machine. This is yet another illustration of how Soviet laboratories are compelled by the condition of the economy to be self-reliant, even self-sufficient. For the production of the superconducting magnets, for example, the institute plans to build a factory with three production lines, each of which should yield one section of the magnet-ring a day. But before that can be done, it will be necessary to hire the 700 skilled workers who will man the plant, but the housing for them and their families will have to be finished first.

The result is that, since the institute was founded in 1963, the population living on the site has grown to 30,000, a third of them children, all of whose needs must be met by the institute's administration. Scientists as such are a small proportion of the total, a mere 500 or so, Three times as many people are engineers. A further complication is that the nuclear research institute at Dubna (now technically an

The new proton accelerator, still at the underground-railway stage.

Thermonuclear fusion helps to build bridges

THE Soviet Union has had a special interest in thermonuclear fusion for civil power production since 1951, when the youthful Andrei Sakharov and the late Igor Tamm sketched out the nuclear physics of fusion and the principles of the magnetic confinement of hot plasma. A group led by the late I.V. Kurchatov (who startled the world by telling the Soviet tale at a meeting in Harwell, England, in 1956), embarked on the construction of an experimental toroidal machine called "*alpha*".

Soviet plasma physicists are especially pleased that Kurchatov's talk at Harwell launched a programme of collaboration that has survived the political upheavals since. Speaking of a mutual friend, one said "he stayed in my flat one summer with his family". For the Soviet Union, the momentum generated in the 1950s is undiminished; now there are six centres in a programme coordinated by academy vice-president Velikhov.

Novosibirsk is one of them, and a monument to the enthusiasm of the late Andrei Budker for the doctrine that particle physics machinery need not be expensive. Budker migrated from Kurchatov's Moscow institute in 1958 together with forty colleagues, in the vanguard of the response to Khruschev's enthusiasm for the opening of Siberia.

The Institute of Nuclear Physics continues Budker's two lines of interest — a family of electron accelerators of which the best-known 5.5 GeV machine (damaged by fire in 1985, now being refurbished and upgraded) and thermonuclear experiments using the magnetic mirror principle in which plasma is confined in a restricted region by a 'magnetic bottle'.

The director of the plasma physics division at the institute, Professor Dmitri Rotov, who worked with Budker for ten years, says he was "not a simple person", which is a way of saying that his enthusiasms would sometimes get the better of other people's feelings.

The result, however, is a collection of ingenious machines for manipulating plasmas as well as electron beams. (The nuclear physics institute also manufactures a range of electron accelerators for sale to Soviet and overseas industry.) The latest magnetic mirror device is a machine with 20 tesla fields at the mirror ends which is compact enough for its operators to hope to beat the plasma instabilities by more and more precise machining.

So will magnetic mirrors be the answer? Rotov echoes those elsewhere who say that much will hang on the success of the large tokamak machines now being built. The Kurchatov Institute in Moscow is building a machine, due to be operating a year from now, intended to go beyond the performance of the Joint European Torus, sited in Britain, that will in roughly five years determine the immediate fate of other routes to fusion. Again as elsewhere, the magnetic mirror people have a sense of urgency.

The plasma physics division of the Physical Technical Institute at Leningrad (not a university, like its Moscow namesake), is working nearer to the ground, on the development of diagnostic techniques for toroidal machines against the commissioning of the new machine at the Kurchatov Institute next year. Professor V.E. Golant, the director, is especially pleased with the the use of Compton scattering for the measurement of electron temperature and of resonance scattering (excited by dye lasers) for the measurement of neutral particles.

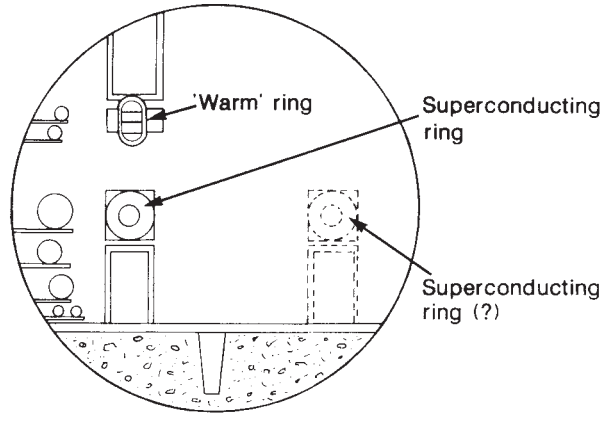
Like others in this business, Golant is delighted with the intimacy of international collaboration in the fusion field which, he says, allows the people in his division a total of some 6 man-months a year of time at fusion laboratories elsewhere. It is almost as if the effort will have been worthwhile on that account even if thermonuclear fusion proves not to be economic. □

international institute run by a consortium of socialist countries), uses roughly 30 per cent of the beam time at Serpukhov, so that the substantial contingent from there must also be accommodated on the site.

What is it all for? Naturally, Serpukhov has no self-doubt, but is sufficiently persuaded of its achievements since the 70 GeV machine began operating in 1967. In the past few years, for example, it claims to have made the running in the exploration of 'glueballs', the electrically neutral entities consisting only of the gluons that hold quarks together and which are left behind when, for example, the quarks of which hadrons are made are removed, perhaps by their decay.

At the end of last month, there was mild excitement that the team at Serpukhov had gathered evidence for the glueball (with a mass equivalent of 1·8 GeV) that decays into four neutral pions (each of which is recognizable by its further decay into two gamma rays).

Serpukhov's role as a stimulant of Soviet industry is harder to define. By 1993 (when the first of the superconducting rings is planned to be in operation), the institute will be operating the largest cryogenic plant in the Soviet Union (generating 30,000 litres of liquid helium an hour), which will be a tangible demonstration of what can be accomplished at 4·2 K. But the Efremov Research Institute for Electrophysical Instrumentation at Leningrad, which manufactures the components such as the warm magnets which are not so specialized that they must be made on site, is said to have more direct links with the electronics industry.



A cross-section of the tunnel for the proposed proton–proton colliding-beam accelerator at Serpukhov, showing the placing of the normal-temperature (warm) magnet ring and the first rings. To the right is the space allowed for the second superconducting ring if permission for that should be forthcoming in the mid-1900s.

Perhaps the most durable if least tangible by-product of the institute's existence lies in its international relations, which are extensive. French groups from Saclay appear to have been the first to visit Serpukhov. Since then, there have been close on 50 collaborations with groups in the United States and at CERN. Only last month, the institute held a week-long workshop to talk over the experiments planned for the new machine. What these mean in terms of physics is anybody's guess; for the time being, people are simply planning to build detectors almost certain to be useful.

Behind last month's workshop lies the hope that it will be possible to interest potential partners in constructive collaboration. The organizers boast that even Japan was represented at the workshop. Nearby Dubna, now constituted as an international organization run by a socialist-bloc consortium, is a way of drawing in the rest of the world.

What will be the response? The hard professionals appear to recognize that everything will depend on their capacity to build a reliable new machine, and on time. The 70 GeV machine was a crucial couple of years late. □

Research institutes

Work, live and die together

To adapt the first sentence of *Anna Karenina*, successful research institutes are all alike, but unsuccessful institutes are unsuccessful in their own distinctive ways. The Soviet Union overflows with institutes in both categories.

The research institute is the indivisible unit of Soviet science, not merely within the Academy of Sciences of the USSR but within the corresponding Academies of Medical and Agricultural Sciences, the separate academies of sciences of the fifteen republics that make up the Soviet Union, and the host of production ministries. By some counts, there are more than 4,000 of these establishments altogether. Defence research is something else again.

The Soviet Union does not of course enjoy a monopoly of research institutes, which have been a part of the landscape at least since the foundation of the Rothamsted Experimental Station in Britain in 1843. In most countries with research pretensions, there are usually traces of a running battle between those who say that institutes are the most effective agents of research and those who complain that research should not be separated from the education of students. The result is usually a quantitative compromise: some institute research and some based in university departments.

The Soviet Union avoids this argument by a compromise: for the sake of effectiveness, let institutes be dominant in research, but make them responsible for a substantial part of the education of the young. Commonly, senior institute people also hold appointments as university professors, thus influencing the curriculum directly. In return, they become personally responsible for the education of some students from the end of the third year of the first-degree course. But the responsibility is also a benefit — they can select the better students as research students, candidates and, eventually, as colleagues. In research, the universities draw the short straw.

Another distinctive feature of Soviet research institutes is their intensely personal character. Many institutes are stamped by their founders with a distinctive style. Peter Kapitza's Institute of Physical Problems in Moscow is renowned for the way in which people are allowed to follow their own instincts. They call it the "Cambridge model". Academician Yuri Osipyan, director of the Moscow solid-state physics institute, claims that his own laboratory is run on the same lines — and, indeed, there are people who have radically changed their fields of research (see p. 793).

Sometimes, a powerful director's idiosyncracies may long survive him. Although Andrei Budker, the founding director of the high-energy physics institute at Novosibirsk, has been dead for ten years, tne laboratory still hallows his rule that a person should not be away from the institute for more than four weeks at a time, an inconvenience to say the least when a complicated collaboration with a distant laboratory may be necessary.

It is more common that a director rules, or seeks to rule, with a rod of iron. The institute has a mission, which the director must carry out. To this end, he (the gender is deliberate) has several important weapons at his disposal: the right to sanction business trips within the Soviet Union, the right to allocate resources and the power of preferment, especially the appointment of laboratory heads.

The appointment of a director is the responsibility of the appropriate disciplinary division of the academy, but appointments must be ratified by the praesidium. It seems to have been common that a director is replaced at death by his faithful right-hand man, often the person to whom a vigorous director would have delegated day-to-day administrative responsibilities.

Appointments, called elections, are for renewable periods of five years, as — technically — are all appointments at academy institutes. For junior people, the requirement is a formality; senior people say "What can we do, when we know there is nowhere they can go?". But the five-year renewal rule can be an effective way of forcing out more senior people, even directors (whose renewal must be decided by the appropriate academy division).

In the internal management of an institute, the director works with a scientific council, of which he is *ex officio* chairman. Technically, the council provides advice both to the director and to the academy. The extent to which directors consult their scientific councils on crucial issues such as the allocation of resources seems to vary from one institute to another. Some institutes maintain a second scientific council to administer the award of research degrees (which must be accredited by a national supervisory board).

The general understanding seems to be that a confident director need not await the advice of his council on important policy issues, but that a director too often in conflict with his council may well find that his appointment is not renewed.

An institute's laboratories are its operational units. Again, their effectiveness depends on the effectiveness of their heads. Some laboratory chiefs appear to consider it their public duty to shoulder administrative burdens so that younger people may get on with more creative work. Others are active scientists functioning as do leaders of research groups in the West; they resent the fact that the practice of science entails such vast amounts of paperwork in the Soviet Union. (Their US equivalents complain that all their time is spent in making research grant applications, but that has at least the appearance of being a creative activity.)

It seems much more common in the Soviet Union than in, say, the United States for junior people to be told what to do (and by when) by the heads of their laboratories, who frequently also exercise the '*droit de seigneur*' of adding their names to the articles written by members of their staffs.

Personal considerations such as these are further intensified by the immobility of the Soviet profession. For all but the most senior people, those likely to be called to higher things, there is a high chance that the institutes people join after the first three years of their undergraduate courses will be those at which they spend the whole of their professional lives. If they find themselves in one of the institute towns that have grown up during the past thirty years, they may well have have the

Fate's fatalities

INSTITUTE directors are usually powerful fellows, but they are also occasionally more vulnerable than their subordinates. Recent events at the Moscow Institute of General Genetics, founded in Moscow after the fall of Lysenko, are an illustration. Until three years ago, the director was Academician N.P.Dubinin, a legendary anti-Lysenko figure who moved to Moscow after being forced out of the Institute of Genetics at Novosibirsk (see p. 800). At the end of his latest five-year stint as director at Moscow, his appointment was not renewed.

Dubinin appears nevertheless to have been able *in absentia* to make life uncomfortable for his well-liked successor, one A.A. Sozinov (not a Lysenko adherent either), who has taken himself off to Kiev, leaving the institute's staff up in arms about the succession. They are particularly incensed that Sozinov should be succeeded by a protégé of Dubinin's, now president of the Moldavian Academy of Sciences.

So, earlier this month, the institute had taken democratization into its own hands, seeking to choose a new director from more than sixteen names by voting on the matter. The hope is that the academy will think it prudent to accept the institute's choice, but it is far from clear how the matter will be resolved.

Sagas such as this are uncommon but not rare. It may even be just that the people at the top should be the more at risk. But it needs very little imagination to appreciate how a row like this can bring people's output of research to a dead stop for months on end. □

same people as close neighbours during the whole of that time (whence some of the attraction of Moscow and the other large cities).

It is asking more of flesh and blood (not to mention mind) than they can bear to expect that everyone can live happily, and remain creative, in such circumstances. Small disputes grow into entrenched quarrels with the passage of time, just as friendly competition may be magnified into envious rivalry as success turns out to be distributed unevenly. Playing laboratory politics may become not merely a sport but a means of survival. Especially in the smaller institute towns, the determination to survive is fortified by the knowledge that finding another job in the locality will be difficult.

The result, inevitably, is that institutes carry on their books many people who have long since gone to sleep. Quantitative estimates are of course impossible, but institute directors estimate the proportion of sleepers at their own laboratories at between 10 and 30 per cent, sometimes venturing ominously that it may be much greater elsewhere.

The giveaway here is not the great size of many Soviet institutes, which is often a consequence of their need to be self-contained enough to make their own research materials, but what people at all except the highest level in the hierarchy have said in open conversation; they name the people, describe how they fill their days, protest that many of them practise "false science" (a word for the kind of work that could not be defended at an open colloquium) and, charitably, add that they sympathize with the reasons why those named cannot be dismissed. There is nowhere for them to go, or to live when they get there.

So, if you are from the West, think of the Soviet research institute as if it were a ship sailing on an endless ocean. Its complement of passengers changes when some occasionally get off and others join. From time to time, some passengers even learn to sail the ship. Everybody knows the destination ("Make thermonuclear fusion economic", or something like that), and knows how far away it is. How do you spend your time? On science, or at beating the other fellow at bridge?

The academy will be deeply hurt at this complaint, which it will say is subjective (which is true). But the observation derives from a great many conversations with people at a score of academy institutes. Some institutes are in good shape; others seem to dissipate their collective energy on battles within their walls. That is not surprising; the same would happen whenever people were bound together so vicariously, and for so long, whence the reference to *Anna Karenina*. But the practice seems peculiar to the Soviet Union and to China. □

Must the young carry all the weight?

THE youngest institute people have the hardest time. Typically, they will have first entered their institute after their third undergraduate year and will have done well enough (half do not) to be kept on for the usual three-year stint working for the degree of cadidate of science (the Soviet equivalent of a PhD).

Although a proportion, usually again a half of those graduating at this level (but there is no way of telling in advance what the numbers will be), are sent to the figurative Siberia of a ministry institute.

The lucky ones kept on (who, if male, will be 27 years or older, allowing for two years of military service) must look forward to several years of personal deprivation. Starting salaries are 125 rubles a month, no more than research students' stipends in fields such as mathematics and physics. But the next step on the academic ladder, the DSc by thesis which allows a person to qualify as a professor of an associated university, is not usually reached until a person's late thirties or early forties.

In the circumstances, it is mystifying that younger members of institutes do as much good work as is apparent. Marriage becomes an economic and social necessity (for the sake of a second, often greater, salary and the chance of an apartment rather than a hostel bed).

But at this stage, children can be a great misfortune. One married couple, with appointments at different institutes and now saddled with a child for whom they must care in a one-room apartment, take turn and turn about at showing up at work, one in the morning and one in the afternoon.

A young person's advancement in these early years depends crucially on his or her laboratory chief, many of whom are able to bring out the best in junior colleagues. But laboratory chiefs sometimes take against the people they have chosen a few years earlier, assigning them the most tedious tasks or keeping them on low salaries. Most people at this stage of their careers take in translation or abstracting work if they can; teaching, the rates for which have recently been doubled, can bring in more than 10 rubles for a two-hour stint (but is usually in the gift of the laboratory chief).

Working women have the rough edge of the marriage bargain. Their husbands, even if they share the view that children would be a misfortune, still expect them to shop for food, a matter of standing endlessly in line, often during working hours.

The climate is something else again, in most of the Soviet Union. So too is the lack of personal (as distinct from public) transport. Institute people are lucky about vacations, with up to six weeks in a year (but a couple of weeks for two at the Black Sea can cost 1,000 rubles). The mystery is not that they do so little but, in the circumstances, so much. □

Shemiakin Institute

Biotechnology's palace in the sun

THOSE who believe all Soviet researchers to be ill-equipped should disabuse themselves by visiting the Shemiakin Institute of Bio-Organic Chemistry, in an inner Moscow suburb. The institute is identifiable by the sculpture of the valinomycin molecule alongside the entrance to the administration block, one of four seven-storey hollow cubes joined together at their corners.

The buildings and its contents have cost $150 million in 10 years, more than half of that in 'currency'. The money shows. The building itself is a startling departure from the starkness of Moscow's monolithic apartment and office blocks. One academician complained at a meeting earlier this year that the sculpture cost more than the research it commemorates.

The interior, by Finnish and Yugoslav contractors, is luxuriously appointed; the floors of the administration block not covered in wall-to-wall carpeting are simple marble. Elsewhere, there are laboratory corridors with red Asian runners down the centre. The council meeting-room, dark wood with tables to match, would make a boardroom for a multinational from the West. The principal auditorium, with over-stuffed tip-back seats in apple green, might be a luxury cinema were the seats less steeply raked.

For the rest, the institute building has a computer centre, a series of pilot plants, a block in which experiments with radioactivity are concentrated and a health-care centre for the staff. Working laboratories are neat and well-constructed, usually in pairs with connecting offices against the external walls (and cold rooms and the like in the dead spaces). Much of the equipment carries labels that should be familiar from the advertisement pages of *Nature*. An internal greenhouse in a connecting internal corridor doubles as a phytotron and decoration.

Physically, the laboratories are better than the best in the West, but emptier. Genentech's workforce in San Francisco, housed in what is by comparison a rabbit-hutch, might go on strike after a sight of the new Shemiakin Institute.

That is probably now misnamed, although many of the 220 qualified people

on the site (there are 1,200 altogether) are chemists by training. For the institute, originally founded as a natural product laboratory in 1959 (Lysenko's time), is now the spearhead of the academy's effort in genetic engineering. It is the creation of 52-year-old Academician Yuri Ovchinnikov, vice-president of the academy and a long-standing enthusiast for biotechnology. (Ovchinnikov, who has been a frequent visitor to the West in recent years, is now gravely ill.)

Just why the building should be so lavishly equipped is far from clear. Is it meant as a model of excellence in laboratory furnishing at which others may strive? A recognition of the importance attached to genetic engineering in the Soviet Union? A mixture of the two? Or something else?

Whatever the case, students, of whom there are 50–60 a year, are dealt with well. They have an excellent set of seminar rooms and laboratories (plus Asian corridor-carpets). The institute organizes subdepartments within biology at the Moscow Physical-Technical Institute and at the Moscow State University.

What of the research? Acting director V.F. Bystrov, himself a structural organic chemist, says that a wide-ranging programme of research has taken the institute more or less to where people are in the West. The production of monoclonal antibodies is routine, there are diagnostic tests for infection by the viruses of AIDS (acquired immune deficiency syndrome), hepatitis B and, interestingly, potato mosaic virus.

Yuri Ovchinnikov — well known in the West and now author of a major textbook.

The development of a technique for producing molecules of tumour necrosis factor (TNF) is complete, and animal toxicity tests preliminary to clinical trials with people are about to begin; Bystrov seemed slightly startled at the suggestion that TNF may be like yet another lymphokine — a naturally occurring biological material whose effects are somewhat ambivalent when tested *in vitro*.

Ovchinnikov's many detractors (envy of the extravagance is rife) overlook his achievements. He has been an enthusiast of genetic manipulation since 1972. At an early stage, he recognized that the prosecution of research in molecular biology would require a supply of materials such as restriction enzymes and radioactive chemicals; there is now a plant at Riga, in the Baltic Republic of Latvia, which produces restriction enzymes and which is also a subsidiary centre of research, especially in toxicity testing and the organization of clinical trials.

Ovchinnikov has also arranged for the production of radioactive chemicals at Tashkent in Central Asia. It cannot be his fault, but that of a production ministry, that the supply is often interrupted, especially in the summer, for periods longer than the 12-day half-life of phosphorus-32. Arrangements for the production of oligonucleotides of specified sequence have not panned out so well.

The Shemiakin Institute is not just a laboratory, however, but the centre of the Soviet Union's web of interest in biotechnology: Ovchinnikov is also chairman of the 65-strong interministerial committee coordinating the activities of more than 200 scientific institutions with an interest in biotechnology. The institute also runs the academy's internal councils on biological membranes and bioorganic chemistry, publishing the two associated Russian-language journals. Ovchinnikov's limitless energy has nevertheless enabled him also to write a major textbook (*Bioorganic Chemistry*, Prosveshcheniye, Moscow, 1987) as well as to participate in many scientific papers. □

Lebedev Institute

High-temple of Soviet physics

THE remarkable Lebedev Institute is a geographically puzzling place. Opposite the main entrance is a building that might be mistaken for a Grecian temple were it not wisely enclosed, a few metres behind the Doric columns, against the hard Moscow winter. That turns out to be the administration block.

The director, Academician N.G. Basov, and his colleagues, are to be found some hundreds of metres behind the imposing entrance, in a laboratory block built, as V.L. Ginzburg quickly explains, on a plot of ground on which people used to grow potatoes in the years immediately after the Second World War.

Closer inspection shows that the institute is not nearly as old as it looks. The first building was in reality put up only in 1951. And behind the imposing facades, many of the laboratory buildings are faced with simple brick. The campus is a kind of simulation of MIT's mix of mock-factories and mock-palaces.

The Lebedev Institute, with all its ramifications, is probably one of the most important concentrations of physicists in the world. Los Alamos would probably give it a close run for its money (and would not have as great a proportion of theoreticians on its payroll of 4,000, at least a quarter of those professionally qualified).

Basov was awarded a Nobel Prize for his contributions to the development of lasers, and he remains active in quantum optics. But he and V.L. Ginzburg, the chief of the theoretical physics laboratory, are anxious to make it plain that the reputation of their institute stems from the great men who have passed through it: Mandelstam, Lansberg, Tamm, Landau, Fock, Fraenkel . . .

Ginzburg, almost certainly the most-active septuagenarian physicist now working every day, and musing on the delectable prospect that he might give up his administrative responsibilities next year, suggests that finding a successor may be difficult with the intelligence that "Sakharov has not been back very long".

Basov is more bullish about the Lebedev. He notes that since the foundation of the institute with the arrival of the St Peterburg academy in Moscow in 1934, with its first director Sergei Vavilov (the brother of the geneticist), no fewer than a dozen academy institutes have been created by the centrifugal tendencies of its various members. Dubna, the now international nuclear research centre, is one of them. "The devotion of our people to science", says Basov, is "extraordinary, enormous".

Ginzburg's view is longer and more down-to-earth. "Just after the war, physics was at the top of its peak — people respected physics more than they respected posts [a reference to the late Lev Landau]". But then salaries went up, there were more people, then other people's salaries increased, now there are probably too many people . . .

This is not the jaundiced comment it may seem. Ginzburg is as active as ever, but the Lebedev is sadly hampered by the serious lack of up-to-date equipment. It would like to be more like Bell-AT&T Laboratories than it is (and is ready to show off a compact electron tube 10 cm across with a spot-size of 10 μm across that may be able to drive a cinema screen through a suitably laser-etched mask). But for lack of equipment, the theoretical division grows at the expense of the others. Ginzburg and Basov seem to have teased each other about that often. □

Soviet universities

The largest system anywhere

By any yardstick, Soviet higher education is gargantuan. The student population is roughly 5 million, with more than a million students recruited every year. These people are educated in a total of nearly 1,000 institutions of various kinds (896 to be precise), most of them with a vocational bent (towards medicine, management or foreign languages, for example). There are 69 institutions known as universities on account of the catholic range of courses on offer.

The standard curriculum requires that a student should take five years over a first degree (but institutions such as the Moscow Physical-Technical Institute allow their undergraduates six years). The first three years of the first-degree course are spent within the walls of the institution concerned, but afterwards a student will be transferred physically to an establishment of a different kind (from a university science department to a research institute, for example, or from a pedagogical institute to a school).

Entry to higher education is competitive, and based on performance in a national school-leaving examination, but some universities run their own competitive examinations to select from among otherwise qualified applicants. For the Soviet Union as a whole, and for institutions of all kinds, the proportion of high-school leavers going on to higher education is just under 20 per cent.

It is a sign of the times in the Soviet Union that the University of Leningrad is planning to advance the date of its own examination from August to July each year so as to be able to compete more effectively for the able students at present creamed off by the State University of Moscow, whose examination has traditionally been in July.

The competition for able students, especially among the universities, appears to have intensified in recent years, partly as institutions have come to recognize that their long-term future will be largely determined by the eagerness with which their graduates are welcomed into the labour-force. The result is that, as in the beleagured British university system, universities have taken to persuading their teachers to excite the interest of potential applicants by delivering popular lectures on some aspect of their science.

Universities and polytechnics (as distinct from monotechnics), being the chief source of recruits for the research institutes, necessarily have close ties with members of institute staffs, many of whom will be nominated as professors and who may, in that role, be responsible both for organizing the whole of the teaching in their special fields and for providing the teachers, usually from among immediate subordinates at their institutes.

The practice seems to vary from place to place, but there are at least some places (the University of Novosibirsk, for example, see page 795 where particular institutes are responsible not merely for education during the closing phase of a first degree but for the first three years as well. First degrees are awarded by universities themselves, but most research degrees are the responsibility of scientific councils of research institutes subject to the approval of national accreditation commissions.

The exceptions are the research degrees awarded directly by those universities that have been able to retain their own research institutes in particular fields. The Moscow State University, even by the consent of its rivals, is the outstanding multivalent university of this kind, but Leningrad (page 796) is also proud of having kept some research capacity within its walls.

Either way, the award of a research degree is a fearsome business. Both Candidates' (PhD) degrees and DScs are awarded on the basis of a thesis (usually a published monograph in the second case). The awarding authority is obliged to circulate copies of the thesis to people knowledgeable in the field, each of whom is entitled to submit critical comments for consideration by the examining board. (Many do, by all accounts.) Finally, the thesis must be defended in public by the applicant, with or without assistance from members of the audience, any of whom is entitled to intervene.

Soviet universities appear to be more like educational mills, if also less confusing, than their counterparts in the West. Part of the explanation is that departments deeply penetrated by the external research institutes with which they are associated are necessarily somewhat isolated from others. Where the curriculum is dominated by a specialized institute, the pattern of teaching also appears inflexible. Why should a bunch of plasma physicists be sensitive to the needs of biologists, for example?

The isolation of vocational education, for example medicine, in separate institutions also deprives Soviet universities of the multi-faceted character called richness in the West. If universities (other than Moscow and Leningrad) were more directly engaged in research, there would be serious difficulties about interdisciplinary coordination; as things are, coordination must be sought through the cumbersome high-level committees involving academies or even ministries.

The Minister of Higher Education, who has contentious problems on his immediate agenda, seems content with the present structural arrangements; in due course, he may find it worth while to seek a merger of vocational schools with existing universities so as to create more interesting institutions (whose administration might be simpler). □

Rationalization

Hunt for quality amid quantity

Minister of Higher Education Dr G.A. Yagodin, appointed by Mr Gorbachev earlier this year to a crucial and previously neglected post, has saddled himself with an invidious task: he is committed to carrying out an evaluation of the quality of the institutions of higher education throughout the Soviet Union so as to tell which should be encouraged to grow and which to shrink.

Only recently, he says with some pride, his ministry has actually closed an institution — the economics institute (essentially a management school) at Baku in Azerbaijan (see p. 783), opening a branch of the Leningrad Institute of Economics there instead. But he has not yet experienced the full fury of important academic institutions scorned.

Yagodin, a vigorous man in his mid-fifties, speaks English excellently, on the strength of a two-year spell studying chemical engineering at Imperial College, London. He talks of the function of higher education in rounded social terms, not merely those appropriate to the supply and demand for specialized labour, which is understandable in one who was snatched away from the rectorship of a university to his present post.

Yagodin does not regard the prospect of evaluating university quality and performance with quite the trepidation that the recent experience of the University Grants Committee in Britain might suggest. "We have no difficulty about Moscow, Leningrad or Tomsk", he says. Most of his difficulties will arise from dealings with the small vocational schools scattered about the country, but the evaluation could also raise national hackles in various of the republics.

For the rest, the process of evaluation will have to be continuous. He plans to set up a system in which outside inspectors will visit universities observing how undergraduates spend their time, on what, and often exercising their right to examine the students to form an objective estimate of what they have learned.

Yagodin goes out of his way to deny the outsider's impression that the quality of

Soviet high-school education is uniformly high. On the contrary, he says, the quality ranges from the "excellent to the bad", and often for no obvious geographical reason; neighbouring schools may be at opposite ends of the spectrum of quality. Like many other educators, and in an outburst of what he calls "political heresy", Yagodin blames the parents of young people for many of these problems.

What with free education, pensions and other social benefits, he says, "people live well in this environment", with the result that young people are often infected by the intellectual laziness of their parents. But young people's ambitions for themselves are often also affected by the "severe egotism" of their parents.

Yagodin reckons that all societies, just now, have difficulties in balancing "material, social and intellectual goals". In the United States, people work for their second car; in the Soviet Union, they buy books they do not read. He wants young people to appreciate that "spiritual demands" should have the upper hand.

The minister is also unwilling to accept that universities are at a disadvantage compared with academy institutes in their conduct of research. The people are as well qualified, are they not? And, in any case, have not university salaries been increased by an average of 25 per cent?

Meanwhile, with such a good Academy of Sciences, higher education can only "welcome the practice" that students should be taught by academy personnel after the third year of their five-year undergraduate course. Yagodin also welcomes the academy's experiment with competitive grants (see p. 782).

So is everything in the garden rosy? Is not the requirement that all Soviet male students should do two years of military service early in their academic careers something of a drag, for example? Yagodin has two answers, one spontaneous, the other, more considered but no less honest. First, he says, military service for students "is not a good idea" and, with an arms control agreement between the superpowers impending, it may not be necessary. But that raises a "very delicate point".

"Look, the other day a German pilot flew over Red Square, and our military were punished. It is a good thing for both of us that his plane was not shot down. But, nearly thirty years ago, Gary Powers was shot down on his flight over the Soviet Union. There are many subtleties in this question of military service for students."

What Yagodin meant — he said as much — is that there is no way in which even such a powerful Soviet citizen as himself can tell whether Marcel Rüst's landing in an unnamed part of Red Square now known officially as Rüst Square was the teenage jape described at the trial a few weeks ago or another plot by the US Central Intelligence Agency.

So what will happen to universities in the Soviet Union? The case of the economics institute at Baku is meant seriously as a warning to other institutions; incompetent (or corrupt) places will be shut down. Yagodin seems quite serious about that.

For the rest, there are the annual budget negotiations, in which Yagodin's personal position is somewhat eased by the circumstance that industrial contributions towards higher education costs are at present about 20 per cent of the total. Like higher education ministers elsewhere, he would like to see industry pay more, if only as a sign that it values the products of the system. One difficulty that could arise is that a successful reorientation of Soviet industry towards a proper respect for costs and values could make plant managers less willing to contribute.

Yagodin's more distant objective for Soviet higher education is that, "having solved the quantitative problem, we must now solve the qualitative problem". He says that the goal must be the development of every person's intellectual capacity to the full, to which end he wants students to be offered (and to assume) more responsibility for their courses (to which end more systematic procedures for grading and assessment will be required).

This account of the problems of higher education in the Soviet Union is not very different from that which would be given, these days, by most of Yagodin's opposite numbers in Western Europe. (The United States, being richer, is different.) But the distinctive feature of the Soviet system is that, with more than 20 per cent of its labour-force trapped in an agricultural industry that still fails to meet social needs, the Soviet Union is probably over-investing in higher education with an overall participation rate of 20 per cent.

The painless solution, of course, would be to make agriculture efficient; Yagodin is probably glad that that is no part of his problem. For the rest, Yagodin has great faith is what he calls this "*perestroika* business", which he believes will keep eating away at people's minds. A few years from now, he hopes, people will be thinking of even more radical schemes than are now in their minds.

They may have no choice. If this "*perestroika* business" does what is intended, the world's largest system of higher education will be operating in a social and economic environment quite different from that by whose influence it has been formed. Will the vocationally orientated (and often small) institutions retain their independence? Will the present restrictions on where students may seek an education survive? One's impression is that even the enthusiasts for change have not yet begun to anticipate its consequences. □

Professional university

Novosibirsk's academic connection

THE State University of Novosibirsk, sandwiched on one side by the special school for physics and mathematics and, on the other, by the research institutes seeking to employ its graduates, is nevertheless not the production-line people might expect. This, after all, is where Budker, the founder in 1958 of the Nuclear Physics Institute, decreed in 1958 that the introductory mechanics course should begin with special relativity.

Even so, the curriculum is a hard grind. Physics students spend nearly a third of their time on mathematics. During the first three years, laboratory work takes the relatively small amount of 160 hours a year, but computer classes (with a PDP11 machine) use up more than half as much time again.

The organizers of the physics course are rigorous in their expectations of students. Professor Nicolai Dikansky, the university dean, says that only 140 of the annual intake of 220 students can expect to graduate. Students drift off by failure or the sense thereof during the first three years, but the annual class is topped up by the immigration of between 15 and 20 students from elsewhere.

In the first four (of five) academic years, a student will have taken 100 annual courses of various kinds, each the equivalent of 64 hours of instruction. By Dikansky's account, students must work hard (and out of teaching hours) to survive. The prize of success is the opportunity to choose, after 4 years, between one of 15 departments related to physics at Novosibirsk, with a monthly salary that may add up to roughly that of a junior employee at an institute — 150 rubles or so.

The fly in the ointment is the static character of the curriculum. Left to themselves, institutes with influence at the university seek to create students in their own image. Novosibirsk has apparently become concerned that there may be more to physics than mere physics. For some years, Dikansky says he has been advocating courses in biophysics to broaden people's interests, even to help them find interesting jobs. He believes he may be near to winning his campaign.

Like all other universities, Novosibirsk is mildly concerned at the decline of interest in science among intending school-leavers. But like others, it is also tempted to believe that the people who matter, those with a sense of dedication to science, have in any case turned up. □

Leningrad

Cosmopolitan Soviet university

THE University of Leningrad hopes to sign a long-term exchange agreement with the University of California before this month is out, but a few weeks ago it was still wondering what to make of a corruption of the word "eukaryotic" in a translated document.

Leningrad's difficulty is that, while second in the Soviet Union (after Moscow) in academic achievement, its cosmopolitan reputation outstrips its capacity to play host, not to mention its budget. The university is also, just now, grappling with requests from 15 universities overseas to accommodate undergraduates on Russian-language courses.

The objective of the proposed agreement with the University of California is to provide for exchanges each year of up to 15 academics and students in mathematics and the natural sciences. California's list of interested partners is a touching roster of its great and good; they also all seem to want to come to Leningrad.

That is not surprising. The Soviet Union's second city is much smarter (and cleaner) than its first. It has three other things going for it: its setting (the Neva is even broader than the Moskva river), its past (as a tsarist capital and the cradle of the revolution) and its reputation as Russia's gateway to the West.

The intended agreement with the University of California is not unique. Leningrad reckons already to have 27 similar agreements with other universities, although most of them are narrower in scope. Many of these arrangements are with universities in Eastern Europe, but the universities of Hamburg, Osaka and Barcelona are on the list. The university reckons to have had 500 academic visitors from outside the Soviet Union last year, most staying for one or two weeks.

Among Soviet universities, Leningrad stands out for the scale of its in-house research. It was the first Soviet university to set up an internal research institute (in 1920), and now has nine of them. By comparison, Moscow has only three. The institutes stand in relation to teaching exactly as do external research institutes; students may be attached to one of them at the end of the third undergraduate year and may be kept on as research students, eventually earning a candidate's degree.

The scale of the research operation is large. The university reckons that its research institutes employ nearly 2,000 people, perhaps 40 per cent of the total staff of the university (of whom 464 are fully qualified professors). There is some dispute about the achievement of the institutes, which are not nearly as well-equipped as are the academy institutes. The university is proud of its institutes, but the full-time researchers at Leningrad's academy institutes are somewhat disdainful. The balance of mutual suspicion is probably exactly similar to that found at academic centres in the West, the University of Cambridge for example.

Leningrad reckons to have 20,000 students, of whom 13,000 are full-time, most of them from the Russian republic, but 1,500 from other republics of the Soviet Union. Most of these external students are people sent to Leningrad as research students by the research institutes of the non-Russian Soviet republics; both Leningrad and Moscow universities are important providers of research training in this way.

The cost of the operation can be defined only with some difficulty. The annual running budget is just over 50 million rubles, of which industrial sources — production ministries and even factories — provide no less than 20 million rubles. The annual grant from the Ministry of Higher Education is 21 million rubles, but the university gets 9 million rubles more from the government to support its research institutes.

University's view of the Winter Palace.

Out of this, the administration has to pay the salaries of its staff, house its students from places other than Leningrad (and pay them a stipend now running at 75 rubles a month) and support its regular academic work. The university is fortunate in that its fixed obligations (staff salaries and so on) come to less than 75 per cent of its total spending, leaving a reasonable sum for discretionary expenditure on equipment and the like.

Of the 9 million rubles a year spent on the research institutes, for example, 4 million rubles is available for research expenditure. Interestingly, the university has begun, in the past two years, distributing part of the research allocation to the physics institute to those who are successful in an internal competition for grants.

The university's most serious financial problem is maintaining its buildings (of which there are nearly 400, many of them across the Neva from the Winter Palace where the tsars used to live, whose storming marked the turning point in the October Revolution and which now houses the Hermitage art collection). The Ministry of Higher Education pays separate grants for the university's capital spending, on buildings as well as major items such as computers.

Much of this money has been spent, in the past few years, on the development of a second campus 35 km from the main site, downstream on the left bank of the Neva. People at academy institutes in Leningrad say this development is a serious mistake, in that it will further separate the university's academics and researchers from those elsewhere in the city.

Meanwhile, the university is fighting a half-hearted battle for the preservation of the buildings on the site on the Neva which it first occupied in 1819. Ceilings in some of the most elegant buildings in the Soviet Union show patches of damp from leaking roofs.

One sign of strain is the drift from science. As at other Soviet universities, people have remarked on the decline of applications from intending science students (2 applicants for every place, compared with 7 ten years ago), despite the number of Leningrad high-schools with a reputation for teaching mathematics and science well.

The university is not complacently persuaded, as are some elsewhere, that all the young people with a sense of dedication to science are still among the applicants, and has mounted a vigorous programme of providing science talks to high-school students in their last two years as well as a scheme for bringing high-school students onto the campus during vacations. The university hopes that the attention it pays to the education of students will count in its favour.

By most of the criteria that would be applicable in the West, Leningrad is as excellent a university as its external reputation suggests. It takes particular pride in its relations with regional industry and government, for whom the university seeks to act as a source of advice. Its most serious problems are the same as elsewhere in the Soviet Union — the immobility of its staff, the personal hardships of everyday life, the bureaucracy and the circumstance that research equipment is generally uncompetitive. It has much to gain from a continuation of the changed political climate. □

Molecular genetics

Another refuge from Lysenko?

THE Institute of Molecular Genetics, not to be confused with the Institute of Molecular Biology or the Shemyakin Institute of Bio-organic Chemistry, all of them in Moscow, is proud to be a direct descendant of an organization set up during the Lysenko period to keep the torch of rationality alight.

What seems to have happened is that Igor Tamm, the theoretical physicist then head of the physics department at the Lebedev Institute,persuaded I.V Kurchatov, then head of the Institute of Atomic Energy, that he should provide a haven for a group of right-thinking biologists within his establishment. A.P. Aleksandrov, until earlier this year the president of the Soviet Academy, but then Kurchatov's deputy, was also involved.

Survivors from this early period recall the embarrassment that could arise when government officials asked members of what was officially known as the Department of Radiation Biology why there was so little going on in the field in which they were ostensibly interested. In the early years, much of their energy was inevitably spent keeping up with happenings in the West, but even in these early and almost clandestine days, it seems that much original work was carried on.

R.B. Khesin-Luyre (Khesin), one of the founders of the laboratory and still much revered, who died two years ago, seems to have been the first off the mark, in the Soviet Union, in the application of biochemical methods to the study of gene expression. Khesin was posthumously awarded a Lenin Prize last year; his colleagues are still politely offended that *Nature* did not know what Khesin had done (see *Nature* **321**, 6; 1987).

Khesin seems to have played a part in the development of Soviet molecular biology a little like that of Max Delbrück in the years leading up to the discovery of the structure of DNA in 1953. The two-week winter schools he organized for members of the institute and others, in the country outside Moscow, had become an annual tradition. They were stopped, after Khesin's death, when the academy refused to sanction their continuation.

The institute now is small by Soviet standards, with some 350 scientists organized into 14 or 15 laboratories — the smallest units which, in a Soviet research institute, are able to make decisions about such matters as who to hire and the execution of a research programme.

Some laboratories have a service function, as those producing radio-labelled compounds for use in this and other institutes. Others, such as the electron microscope laboratory, work on their own problems but provide services for other laboratories at the institute. Still others reckon to keep up with the competition elsewhere, in fields such as structural analysis. The institute appears not to have succeeded in setting up a line for the synthesis of oligonucleotides, with the result that individual members of the institute make arrangements to produce sequences they need with the help of friends elsewhere.

Khesin's mark is plainly to be seen in the present pattern of work; one laboratory works away still on *Escherichia coli* RNA polymersase and another is concerned with the molecular basis of genome instability (the title of Khesin's last monograph) which has become, at the molecular genetics laboratory, a way of referring to the integration of the μu bacteriophage into the bacterial genome.

Equipment is a constant headache but, in a larger sense, not a problem. Standard laboratory equipment is readily available. Institutes differ greatly in their access to hard currency; the molecular genetics institute must make do without equipment bought in the West (but the structural analysis laboratory has acquired an IBM minicomputer for data-handling and processing). The same laboratory has a collaboration with the Daresbury Laboratory of the UK Science and Engineering Council.

Supplies of radioactively labelled compounds and other reagents leave much to be desired. The academy took the initiative some years ago in founding plants for the production of labelled compounds in Tashkent and for restriction enzymes at Vilnus and at Novosibirsk. There are some snags. Radio-phosphorus nucleotides labelled in the gamma position are less active than they might be, while the general interdiction against the purchase overseas of materials that can in principle be bought domestically is more than a mere nuisance when the Tashkent plant breaks down, or is put out of action for routine maintenance.

The general air of shabbiness is more disconcerting. Bench-tops do not always fit neatly against walls. Mock-terrazzo floor coverings (a kind of linoleum) have shifted here and there in the corridors, exposing the boards beneath. Curved joins between vertical walls and floors have been fashioned from unfinished concrete. There is a general lack of illumination (one of the superficial hallmarks of the Soviet Union). One laboratory head was compelled to climb on a chair to extract a document from his filing system (but the general rule seems to be to keep papers in piles on desks and table-tops).

Like others on the academy network, the institute takes in undergraduate students after the first three years of their courses have been completed successfully. Most of its business is with the Moscow Physical and Technical Institute and the Moscow State University, the former with an unusual six-year pattern of degree courses (most are just five years long). The idea is that students should cut their teeth on professional work before they graduate, but the system is also a convenient way of allowing an institute to tell which students it wishes to keep.

The better students are able to pick the departments to which they are assigned, the others take pot luck. For at least the first year, they are formally taught by members of the institute staff. Thereafter, they take on the flavour of people halfway between laboratory assistants and beginning research students.

The laboratory has one distinction that marks it out — it appears to have been the base from which which has sprung a strictly applied laboratory, an institute belonging to the Ministry of Microbiology. People speak of their departed colleague as if he were lost forever, the kind of exile they will never encounter at regular scientific conferences. But the institute, by repute at least, is doing well.

Novosibirsk's special school

Hot-house for young professionals

THE special school at Novosibirsk was last month playing host to a party of students from the Phillips Exeter Academy and Andover, two well-known private schools in the United States. For the school at Novosibirsk, the objective is to broaden students' cultural experience. Later, there will be a return visit to New Hampshire.

Novosibirsk properly regards itself as in the same league as its US partners. The 550 students at the school, itself one of a dozen in the Soviet Union founded on the elitist doctrine that students who are already outstanding at physics and mathematics would benefit from special instruction, have been selected from the whole area of the Soviet Union east of the Urals on the basis of the annual mathematics and physics 'olympiads' organized at school, regional and national level.

Elitism is not as odd a concept in the Soviet Union as might be supposed; on the principle that it is every citizen's duty to serve the state as best he can, does it not make sense that people with special gifts should have special training to make their future service more effective?

Dr Mikhail Lavrentiev, dean of studies at the school and a physicist, acknowledges that there are hazards in selecting

Small, high-tech back in business?

SMALL high-technology business may have a bright future in the Soviet Union. *Perestroika* so far has changed the economic rules to allow for the creation of cooperative businesses, each of which will be owned equally by those who work for it. Some researchers are jumping at the opportunity.

Most attention hithero has been directed at the cooperative businesses springing up in agriculture and horticulture, often the source of goods sold in the free markets. (Family businesses are also now allowed.) But there are at least two technical fields that offer scope for entrepreneurship — the development of computer software and of diagnostic reagents derived by genetic manipulation.

That restrictions on business activity should have been lifted chiefly in respect of cooperatives is easily understood: one of the communists' objections to capitalism is that it enables individuals to exploit the labour of others. Cooperative businesses share the labour and the proceeds, but hitherto have not been encouraged. Even in the changed climate, people must stick with the jobs at which they are supposed to work, while the arrangements for providing cooperatives with working capital are still rudimentary. Nor does Soviet law makes provision for the indemnity from creditors that follows from incorporation in the West.

The origins of the mini-software houses (there are said to be half a dozen in Moscow) are easily imagined. People simply use the computers at their workplaces, but out of working hours, to develop novel or customized software, which they then sell on to customers with whom they have made a deal. Strictly, the institute or other institution should be told of the arrangement, and may in turn require a share of the proceeds.

The molecular diagnostics cooperative being organized in Moscow is more unusual, involving roughly a score of people already. The objective is to identify and then to clone potential biological reagents, bacterial and viral antigens for example. The customers would include those physicians still engaged in private practice.

At least one potential member of this consortium was in two minds, earlier this month, about an invitation to join this consortium. The meeting he had just attended had left him with the impression that the organizers were "simply concerned with making money". He planned to think the offer over. □

such a small proportion of students. The problem-orientated olympiad system ("If you boil a turkey, why does the skin split in only one direction?", for example) requires that students should change intellectual gear. The school has taken to organizing summer schools to help high-school students appreciate its needs.

The school, of necessity residential, is essentially a finishing school, accounting for one or two years at the end of a student's high-school career. There is a girls' quota ("of course, they're weaker in the olympiads"), although Lavrentiev names some of the young women who are doing well at the university.

The school seems less dismayed than it might be at the decision of a small proportion of its students to quit after a month at Novosibirsk, on the spartan grounds that, whether the decision springs from nostalgia (homesickness) or dismay at being a small fish in a large pond, it shows the defectors to be "unfitted for a scientific career".

The curriculum at the special school is no different from that of the past two years at any Soviet high-school, Lavrentiev says. Students will, for example, be taught calculus, but will not be expected to evaluate $\int e^{-x^2}dx$, for example. The objective is simply to give them a more thorough grounding in their subjects.

The school has been running computer classes since 1972, but has only recently made them generally available. And there remains a great shortage of equipment, with the result that students do not have *ad lib* access to computers. Lavrentiev acknowledges that the school may have to strengthen its teaching of biology.

Links between the school and the rest of Novosibirsk's scientific enterprise are close (although some are mystified that the school appears not to have been sending its students to visit the nuclear physics institute). Some 35 institute researchers have teaching responsibilities at the school, despite the inconvenience caused by their frequent business trips.

Graduates from the school account for 40 per cent of the physics and 25 per cent of the mathematics students at the university on campus. Lavrentiev himself says with pride that his career has taken him through the whole cycle, from school to university to Nuclear Physics Institute and back to the school.

So is the school one to which one would wish to send one's son or daughter? The students themselves, who are as bright as paint, say they enjoy the atmosphere and the competitiveness. One says that she has found it helpful to be given this reminder that she may not be as good a scientist as she had thought. Given that Novosibirsk is a good place at which to practise science (and that Moscow and Leningrad are inaccessible), why not? The arguments in the other direction are merely the spartan life, the loneliness and the isolation, but there is a lot of that in the Soviet Union. □

Publishing

Is opportunity durable?

MIR (which means "world") Publishers, one of the largest operations of its kind anywhere, is adapting to *glasnost* with enthusiasm but uncertainly. Most of the company's business is the translation of 200 technical books from English into Russian and the translation of 300 books a year from Russian into several other languages. Vladimir Kartsev, director of the company, explained earlier this month that *glasnost* has provided opportunities not yet fully grasped.

In the old days, Kartsev says, the company was "told" what books to publish; with the help of 700 employees, it discharged its obligations and earned a surplus of about 2 million rubles a year. Now, the company has to decide for itself what books to publish and can negotiate with the publishers of English-language books on the price it will pay for translation rights. But domestic prices, those at which it may sell its products, are still controlled, so that the company has to find some substitute for what would be called a market in the West.

The solution is a system of "readers' conferences". Potential readers are brought to the park-like headquarters in suburban Moscow to say what kinds of books they would like to read, or would find most useful (reference books and books on computers predominate).

What about the books sold on in the other direction? Mir says that "frankly, we have no direct experience" of the wheeling and dealing that passes for publishing in the international market. But, by now, Kartsev and his colleagues may have learned a little from their first visit to the Frankfurt Book Fair earlier in the month.

Kartsev and his colleagues have a few simple ambitions. They want to be able to price the books they publish more flexibly, usually charging more. They think they could reduce their workforce by a half without loss of production, demanding more from people instead. They would like to make it possible to reduce the time taken for publication (they are in their printers' hands) from three years to something much less. (A Compugraphic desktop publishing system is being played with in a shed in the grounds, perhaps pointing the way to that ambitious goal.

Meanwhile, everybody considers what money Mir might have made had not Pergamon Press secured the rights to publish Landau & Lifshitz in English. But Mir stands to do as well from its decision to publish *Scientific American* in Russian, an operation fronted by Professor Sergei Kapitza and run by Mrs Lidiya Shepeleva.

Lenin's Irkutsk

Regional university chases young

LENIN's generosity towards the University of Irkutsk may also be a millstone around its neck: by making the university a depository library, entitled to receive a copy of every book or journal published in the Soviet Union, he may also have condemned it to the fate of being a librarians' university. Asked last month what the university was most proud of, vice-dean Viktor Isayev mentioned (in this order) the library, the computer centre and science education.

The city would be incomplete without a university. Some of the tree-lined streets have rows of old wooden Russian houses, but the centre is dominated by the headquarters of the regional soviet (parliament) and by an unfinished headquarters of unparalleled hideousness for the regional party.

The university has a curious history in the modern Soviet state. At the turn of the

The Trans-Siberian Railway is no longer the only artery to the developing eastern zone.

century, Irkutsk was a prosperous merchant city that flourished on the trade with China. Soon after the revolution of 1917, V.I. Lenin, no doubt conscious of the city's strategic importance, decreed that it should have a university. But the region turned out to be a hotbed of counter-revolution, not quashed until 1921. The university came into being in the interregnum.

In reality, the university is best known for its historical studies (counted as a part of science in the Soviet Union), oriental studies and philology in particular. But Isayev is also proud that 25 full and corresponding members of the 'big' academy have graduated from Irkutsk.

To face facts, Irkutsk is a worthy regional university. The best potential students in science and mathematics in eastern Siberia are creamed off by the special school at Novosibirsk and tend to continue at the university there.

Like other universities in the Soviet Union, Irkutsk is finding it more difficult than in the old days to recruit students into science and mathematics. Law, journalism and historical studies are, just now, most popular among school-leavers, with six candidates for every vacant place.

Isayev says that the demand for a place in the natural sciences has fallen, in the past 10–12 years, to only a quarter as much. Rural students, he believes, are not nearly well enough prepared for science (for which reason the university does what it can to bridge the gap with introductory courses in the early years).

Irkutsk follows the general pattern of education in the Soviet Union by providing its full-time students with full-time instruction during their first three years, and then placing them in semi-vocational settings, research institutes or pedagogical institutes as the case may be. Half of the 10,000 students on its books are part-timers, people already working at jobs in industry and administration, which is an unusually large proportion.

The university would like to grow, but financial constraints imposed by the university's paymasters, the ministries responsible for the Russian Republics as well as the Ministry of Higher Education, may blight that prospect. There is also some anxiety about the proposal that employers of graduates from Irkutsk should pay the university 3,000 rubles, or 30 per cent of the cost of a five-year education, for every graduate hired.

Meanwhile, the university echoes many others in higher education in its belief that the decline in the numbers of people applying for science places represents merely a weeding out of the feckless from the dedicated. "The demand is from serious students still." Even so, Irkutsk has followed most other universities in the Soviet Union in sending its teachers out to the high-schools of eastern Siberia as proselytizers of its cause. It seems to be learning empirically about the market for higher education. □

Siberia's academic town

Grow old gracefully in Novosibirsk

How does a town full of researchers age? The *akademgorodok* of Novosibirsk, the Soviet Union's largest special-purpose civil research community, is some 35 km north of the Siberian city of Novosibirsk, itself on the Trans-Siberian Railway at the western edge of the Siberian plain and on the northerly flowing Ob (not now to be diverted south).

The academic town, founded in 1957, is a monument to the early days of Khruschev's policy aimed at the settlement and exploitation of Siberia. The policy is still working away, but more gradually than that impetuous man would have liked.

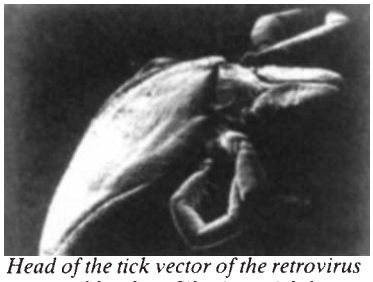
The town has become a major scientific and technological centre. Apart from more than a score of academy institutes, it provides the headquarters of the academy's Siberian Division. There are also institutes of the academies of agricultural and medical sciences as well as of several of the production ministries. There are 70,000 people working in more than 100 institutes. Despite the academic character of much of the work, the deliberate intention is to assist with the application of research to the special problems of Siberia.

Ten years ago, *akademgorodok* seemed an idyllic place, at least in comparison with industrial Novosibirsk, with its tower blocks and empty shops. That remains the case, but the fabric of the town has aged. The birch trees are taller, but the doors and other external wooden fittings of the older buildings plainly show the marks of ten Siberian winters. Some buildings, the special school for example, have grown to look positively slummy. The community's central restaurant has become shabby, dominated after dark by the noise from the standard over-noisy pop group accompanying groups of exceedingly restrained dancers.

What of the people, getting on for 200,000 altogether? They have mostly stayed in their jobs, with the obvious result that the proportion of young children in the community has fallen. More important, some of the people born in *akademgorodok* have grown up not to be scientists like their parents (the town's excellent schools are strong on music, among other things) and have moved away, but often only to Novosibirsk.

For most of the mid-career people in *akademgorodok*, the life is satisfying and self-sufficient. Institute work for many is relieved by a spell of teaching at the town's university or even at the special school (see p. 797), which creams off Siberian high-school students in the last two years of their courses for special instruction in science and mathematics.

For many now working at the institutes, this *akademgorodok* has been the whole of life. Progression from the special school to the university to a post at one of the institutes seems commonplace. Those questioned about the isolation of their lives nevertheless insist that they have no sense of being cut off from the rest of the world.



Head of the tick vector of the retrovirus responsible for Siberian tick-borne encephalitis. The Institute of Cytology and Genetics says it has developed an RNA probe for telling which tick-bodies recovered from bitten people's skin are likely to have been infective, and that hundreds of cases a month are being diagnosed.

Funds for domestic business trips seem ample; in fields such as high-energy physics, visits to laboratories and conferences overseas are frequent. Visitors from elsewhere, even from overseas, are always dropping by.

Even so, there is a mild sense of senescence about the place. Institute staffs are that much older. Administrators are moderately concerned that the pace of recruitment to the institutes is so much less than 30 years ago, when most establishments began from scratch. Some look to expansion off the site, possibly by the creation elsewhere in the Soviet Union of branches of themselves.

The isolation probably also helps to explain the emergence, towards the end of last year, of a curious organization called 'Memory', made possible by *glasnost*, but with explicitly anti-semitic overtones. The organization, in which mathematicians played an important part, is now said to have sunk again without trace.

Might it not in the circumstances have been been better to build the new community closer to the city? This question provoked from the wife of one senior scientist the plaintive cry "Why didn't they put it nearer Moscow?" People can still think of themselves as exiles after several decades in the east.

Part of the trouble is that at least the senior people at the institutes spend much of their time travelling to the capital for committee meetings, sometimes once a week or so. With a sufficiently early start, it is possible to make the return journey within the day. The women tend to be left behind.

So has this best-known of all *akademgorodok*s succeeded? Much depends on the criteria for deciding what constitutes success. That Novosibirsk has grown to become a substantial intellectual centre, comparable with, say, Kiev or Kharkhov, is not disputed.

At least the mathematics and nuclear physics institutes have international reputations for research, while the Institute of Economics and Industrial Engineering has won for itself and for Academician A.G. Aganbegyan, its director from 1966 until 1985, a reputation for economic modelling and, more recently, a considerable influence on the Soviet Union's five-year plans. Now, from a base in Moscow, Aganbegyan is up to his eyes in the planning of Mr Gorbachev's intended economic reforms.

As an educational centre, this *akademgorodok* also seems to have made its mark. The university and its graduates are respected. There is some wry amusement in the town that the university appears to emerge badly from bureaucratic quality assessments, based as these are on counts of the diplomas and degrees awarded to members of its full-time staff. The joke is that most of the teaching is done by more highly qualified members of the institutes.

The economic impact of this concentration of research on Siberia is much harder to assess. Much of the work at Novosibirsk is biased towards applications, especially in fields such as geophysical prospecting and ore-processing so that, to the extent that these sectors of Soviet industry are doing reasonably well, there is likely to have been a worthwhile connection.

To a visitor from the West, however, the most surprising feature of the town is that it is not, by now, the centre of a huge science park filled with spin-off industries. It is true that many of the ministry institutes are in reality pilot plants, while the nuclear physics institute (see p. 790) has used its experise to build up a business in industrial electron accelerators. But can it be too late to encourage the town to become not just a source of research for distant industries, but a centre for much less formal industrial development making direct use of researchers' enthusiasm for their innovations? That way, this *akademgorodok* would seem a little less like an insect preserved in amber. □

Siberian genetics

Putting Lysenko behind them

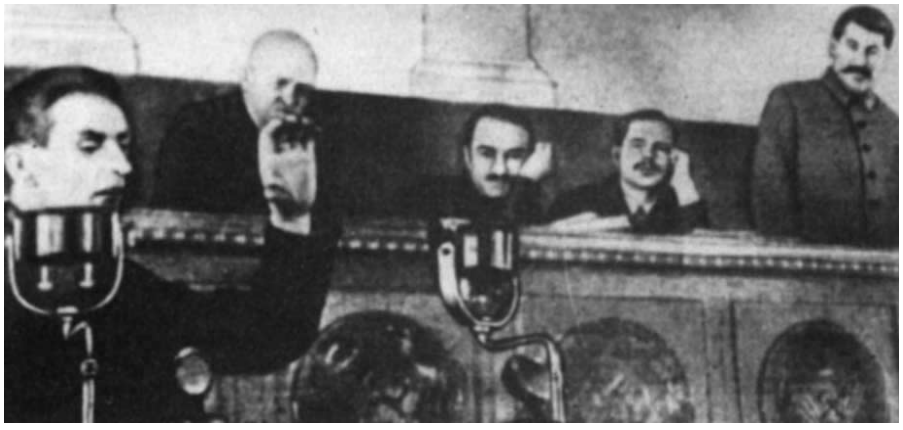
AMONG the many old stratagems for preserving Soviet genetics from Lysenko is the Institute of Cytology and Genetics at Novosibirsk's *akademgorodok*, founded by Academician N.P. Dubinin in 1958. The stratagem did not survive long. Part of the legend of the institute is that Khruschev thought it worthwhile stopping off at Novosibirsk, on a return journey from the United States in 1959, to read the riot act to the dissident geneticists. According to Professor Vladimir K. Shumny, who was there at the time, "we sacrificed Dubinin to preserve the institute itself".

For the succeeding quarter of a century, the director was the late Academician D.K.Belyaev. His successor, Shumny, is a corresponding member of the academy. He is showing off apparently the same coloured mink pelts as were on show exactly a decade ago.

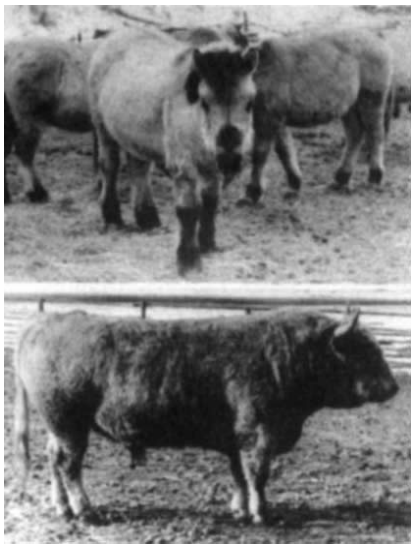
By general agreement, the institute has done invaluable work for the mink industry, chiefly by conventional breeding for arresting colour and other qualities of the fur. It is also proud of its work on the sea buckthorn, cultivated in Siberia for its berries (which are now twice as big as in the old varieties and easier to pick because the bushes have been persuaded to lose their thorns). It collaborates with institutes of the agricultural academy at Novosibirsk on sheep and cereal breeding and, since 1980, has maintained in the Gorno-Altai autonomous region a reserve for 20 aboriginal animal breeds and species, the Yakut horse for example.

As it is now, the institute is a blend of an applied genetics institute and one concerned (for this *akademgorodok*) with being a centre of expertise in molecular biology, which task it shares with the nearby Institute of Bioorganic Chemistry. The work ranges from the maintenance of a nucleotide sequence bank to the transplantation of mammalian oocytes.

Of the total staff of around 700, scientists account for 300, divided among 25 laboratories (including several out-



Lysenko (left) under the proprietorial eye of Stalin at a 1935 rally.



The Novosibirsk geneticists are concentrating on the practical. Like preserving a herd of the ragged Yakut horse (top) and the Yakut cattle, resistant to frost and able to pick up the grass from under a layer of snow.

stations). The institute has latterly been using the techniques of molecular biology in the pursuit of its long-standing interest in gene regulation, wherein lies a more controversial tale.

How to domesticate the silver fox? A successful answer to this problem would have an obvious interest for the fur industry, which would prefer not to have to go hunting for these animals (and people are in any case worried about their survival). This line of thought prompted, 30 years ago, the still-continuing study of the process of domestication, which in turn has provoked the study of the genetics of the inheritance of animal aggressiveness and other behavioural traits.

To vistors from the West, these sound like Lamarckian exercises, but Shumny insists that there is no heresy in his conviction that it is possible to breed for traits such as aggressiveness and its opposites that mark out readiness for domestication. On the contrary, the study founded by Balyaev is said to show that the phenotypic changes accompanying domestication do not imply genomic reconstruction, but changes in the pattern of gene regulation, possibly causing the activation of silent genes or increasing the natural mutation rate. Intracellular influences, such as hormone levels, might serve to switch from one pattern of gene regulation to another.

Why, for an institute that claims to publish a fifth of its research papers in Western journals, has this material made so little impact outside the Soviet Union, or even Novosibirsk? Shumny and his colleagues did not jump at the notion that a review article in *Nature* might do the trick, but one would be welcome, while the refereeing process would be instructive for all concerned. □

New departures

Starting out afresh in Siberia

If being sent to Siberia is a misfortune, electing to go there must be a folly. But there are compensations, which explains why Dr Mikhail Grachev has taken the job as director of the Institute of Limnology at Irkutsk, on the Trans-Siberian Railway at the point at which the Siberian plain joins the mountainous region to the east.

Grachev has persuaded a score of colleagues, some of them at odds with their institute directors, to migrate east with him from Novosibirsk. Last month, he was living with ten of his new colleagues in a communal apartment, but hoping to have been given somewhere more permanent by now. Meanwhile, there are people in Novosibirsk packing their equipment in preparation for slipping into newly emptied beds.

The venture is important both as a test of the flexibility of the administrative structure that seems normally to inhibit the movement of people and as a test of *glasnost* in the narrow sense of interdisciplinary communications.

Irkutsk is not everybody's vision of Siberia, but a typical small Soviet city, with close on a million people, a university and the regulation Intourist hotel. Across the river from the old city is the *akademgorodok* (for "academic town"). This is where Grachev has settled at the institute of limnology. His brief, from the Siberian Division of the academy, is to make Lake Baikal the centrepiece of a cross-disciplinary research programme whose guiding principle is merely that the lake is extremely interesting. One long-term but incidental effect may be to sharpen the tone of Irkutsk's *akademgorodok*. In a city five hours ahead of Moscow and an hour behind Tokyo, it is a little like being asked in eighteenth-century England to go and manage a colony in the New World: remoteness gives you independence, but also great responsibility.

Grachev, a soft-spoken but sharp-tongued man in his early forties, seems undismayed. An ecologist himself, Grachev has already gathered together a hard core of people who know about aquatic microorganisms, and has also formed a group of dedicated instrument-builders whose techniques may be applied to the study of Lake Baikal while providing grist for the mill of the always-nascent Soviet instrumentation industry.

One of his colleagues is Sergei Kuzmin, who has developed a liquid chromatography instrument, with a built-in double-beam spectrophotometer which, he claims, requires less of the expensive purified solvent than similar machines available in the West. (The group is still resentful that the Swedish company LKB decided some years ago not to market the device internationally, noting that more than 1,000 copies of it have so far been manufactured for domestic use.)

Another zealous designer is working with a Leningrad institute on the development of a novel and sensitive molecular mass spectrometer (intended to be used for baseline measurements of the present contamination of Lake Baikal) as well as a group skilled at synthesizing oligonucleotides (see pp. 802 and 792 respectively). Partly because of the second of these developments, Grachev is looking for molecular biologists keen to apply their techniques to the evolutionary history of the species of Lake Baikal.

There are more radical ideas in the air. Grachev takes it as a given that Lake Baikal is too interesting to be exclusively a Soviet preserve. He wants to make it international. Ingeniously following the *glasnost* principle that "what is not forbidden must be allowable", he has concluded that there is nothing in the academy's rules to prevent him employing specialists from abroad (a point confirmed in Moscow by a slightly startled Mr I.A. Timofeev, a member of the academy's Department of Foreign Relations). To begin with, Grachev would like to find a biologist from the West willing to make an "excursion" to Baikal and then perhaps to send a research student prepared to live on "our rubles" and to "share our life".

Part of the hidden agenda is to make a stronger link with science in the West. (Interested readers should write directly to Dr Grachev at the Institute of Limnology, Far Eastern Branch, Siberian Division of the Academy of Sciences of the USSR, Irkutsk, but *Nature* would also like to know of any expressions of interest.

More immediately, Grachev has designs on the University of Irkutsk. "We should be doing more teaching", he says. In due course, no doubt, that will be the case. But it will be an uphill task. So, too, will be that of sustaining the spirits of the new team through the hardships of the winter now beginning.

Grachev's adventure, like those of many other Siberian pioneers, is an intellectual risk. Distance from Moscow may offer opportunity, but also the chance of being forgotten. Grachev believes he has the personal backing of Academician Valentin Koptyug, president of the Siberian Division, but to the extent that his success will turn on people's willingness to break with the traditions of immobility and many other backward conventions of Soviet science, even those who have no inclination to join the venture might send him and his colleagues a greeting at the coming winter solstice. □

Environment

Baikal a symbol of Soviet intent

THE most vivid illustration of the scale of Lake Baikal is the sight of the river Angara at Irkutsk, 60 km away from the point at which it is the only exit from this inland sea. The river is more than a kilometre wide at this point, and fast enough to cause trouble for the ferries that ply between the banks.

The lake itself is almost incomprehensibly huge. Standing on the Western shore on a clear day in late September with good visibility, one could see the peaks of the wall of rock rising from the eastern shore but not the roots, still hidden in the surface mist. At the exit of the Angara, the lake is more than 20 km across. Elsewhere, it is wider. The other two dimensions of Baikal are even more puzzling. The length, as any map will show, is more than 700 km, but the depth reaches to 1,640 m, making the lake the largest single body of fresh water on the surface of the Earth. The local estimate is that Baikal contains between a quarter and a third of the world's fresh water. The average flow of the Angara is merely enough to empty the lake once in 400 years.

The structure is evidently a rift valley. Chains of islands on transects of the lake mark the places where earlier mountain ranges have sunk with the rift floor, helping to divide the water circulation of the lake into three largely separate cells. On the face of things, the rift will break out to the sea on the shore of the Sea of Okhutsk, opposite the Kamchatka peninsula.

Estimates of Baikal's age, based largely on palaeontological data, are about 20 million years, but so little is known of the underlying tectonic processes that Baikal must be a happy hunting ground for geophysicists for many years to come.

Biologists' interests may be even more pointed, according to the newly appointed director of the Institute of Limnology at Irkutsk, Dr Mikhail Grachev. The baby seal at the University of Irkutsk's research station up the lake poses a problem: how does it come about than an inland lake supports freshwater adaptations of seawater mammals? There are 2,500 indigenous species in the lake, including bottom-living fishes adapted to the pressure of 1,600 m of water and sponges not very different from those best known in tropical seas. How did this diversity arise in a mere 20 million years?

But none of this is evident enough to give the lake its sense of magic. The approach to the place is more evocative. The road from Irkutsk, winding up the right bank of the Angara, first reaches a promontory 100 m above the sill over which the lake is emptied.

A kilometre further on, there is a hamlet of log-built Russian houses, with their curiously projecting dovetail joints at the corners (and some modern equivalents being built of sawn 25-cm timbers with corners that look the same) together with a working Orthodox church half-filled, on a Sunday evening, with local women. The lakeside road ends a further kilometre beyond, at a dock with a handful of research vessels. Beyond that, there is nothing to scar the shoreline, so that the steep cliffs, painted red and yellow as well as green for autumn, fall directly into the lake.

Communications around Lake Baikal are by the surface of the lake. In the summer, people use boats or even ships, in the winter, sledges or skis. A few kilometres north of the dock, there is a solar telescope operated by the Institute of Astrophysics at Irkutsk perched on a peak several hundred metres above the lake. Twenty km north of that is a research station belonging to the biology department of the University of Irkutsk.

Baikal — a place of superlatives and extremes.

In a cove marked by a series of shoreline pinnacles to the north, there is a comfortable wooden house and half a dozen uncomfortable huts, in one of which a research student is huddled over jars of larvae. A juvenile seal flops about in a shallow concrete pool, ostensibly in the cause of a test of visual cognition of which the best to be said is that it is uncontrolled.

Water quality has made Baikal a contemporary legend. You can tell how pure it is by looking at the pebbles 5 m down, at the edges of the lake. Grachev's group plans to use the most sensitive instruments they can find (or build) to establish a baseline with which later measurements can be compared. The purity of the water explains why the Ministry of Paper and Pulp Industries decided, more than a decade ago, to build two paper-pulp plants on the south-western bank, one on a tributary and the other on the lake itself. There has been moderate uproar ever since. At one stage, the late Peter Kapitza intervened in support of those who fear that the output of 200,000 m^3 a day of effluent from each of the two plants will eventually spoil the lake. Two years ago, the ministry agreed that both plants should be closed, but not until the labour-force moved from Western Russia to operate the plants had been retrained for making furniture.

People say that they were first alerted to the polluting potential of the pulp plants by the sophistication of the equipment for the treatment of the effluent. Quite what damage has been done so far is not yet clear, but a visiting Finnish group this summer is said to have evidence that dioxin is formed during the burning of chlorolignin practised at the two plants.

The saga of Baikal marks a turning point in Soviet opinions, especially the intelligensia's opinions, of the environment and its protection. In due course, the movement will no doubt be recognized to have been part of an indirect protest at the intellectual bankruptcy of the Brezhnev era, and which has now made environmental integrity a goal akin to peace (by means of the abolition of nuclear weapons).

People talk of "our global heritage" and "our common problem", inviting all shoulders to help propel the same wheel. The same opinion-makers are oddly indifferent to the way that Brezhnev's (and even Khruschev's) suburban factories continue to pollute the air their people breathe.

But what about the pollution of Baikal's four-century stock of discharge for the Angara? Early on a sunlit September day, only the people with the sharpest vision claimed to spot the discharge from the chimney stacks the other side of the lake. But as the temperature fell, there was a curious phenomenon: one of several valleys on the other side began to fill with mist, a consequence, the knowledgeable said, of diffusive turbulence stimulated by the distant chimney stacks. Then two dark plumes of smoke appeared silhouetted against the mist, illuminated for an endless half-hour by the setting pre-winter sun. The sailing party, atheists to a man, went off to hear the singing at the Orthodox church. □

Gondwanaland's seasonal cycle

Thomas J. Crowley*†, John G. Mengel‡ & David A. Short*

* Laboratory for Atmospheres, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
‡ Applied Research Corporation, Landover, Maryland 20785, USA

A two-dimensional energy balance climate model has been used to simulate the seasonal temperature cycle on a supercontinent-sized land mass. Experiments with idealized and realistic geography indicate that the land-sea configuration in high latitudes exerts a strong influence on the magnitude of summer warming. These simulations provide significant insight into the evolution of climate during the Palaeozoic, and raise questions about the presumed pre-eminent role of carbon dioxide in explaining long-term climate change.

ON a timescale of tens of millions of years, the most important climate question concerns factors responsible for the maintenance of an ice-free Earth[1,2]. In the past few years, there has been a shift in emphasis away from palaeogeography to CO_2 as the most important variable responsible for such a large climate change[3-5]. But these studies have not adequately addressed the role of the seasonal cycle, which is strongly affected by the land-sea distribution (see, for example, ref. 6, Fig. 7.20), and which has been shown to be important in the climate fluctuations of the past 100 Myr.

In this study, we employ a two-dimensional energy balance model (EBM) to simulate the seasonal cycle on a very large land mass (the Palaeozoic supercontinent of Gondwanaland). We find that summer temperatures can exceed 20 °C at the South Pole if it is in the centre of a large land mass. If the South Pole is at the edge of the land mass, summer temperatures may not get significantly above freezing, even if the solar constant (or CO_2 content of the atmosphere) is increased. This is because the large heat capacity of water tends to suppress the magnitude of warming on adjacent land areas. These simulations may resolve some significant paradoxes of Palaeozoic climate, and raise questions about the presumed pre-eminent role of CO_2 in climate change on very long timescales[8,9].

Resolution of the seasonal cycle is important because it allows us to focus on a critical boundary condition necessary for initiation of glaciation—the summer temperature. It is not possible to initiate ice cap formation if it gets too hot in the summer. We assume that the summer temperature is more important than moisture supply for glacial initiation[9]; this assumption is supported by the simple calculation that, even if total annual precipitation were as low as 10 cm and summer temperatures did not rise above freezing, it would be possible to build a 1,000-m-thick ice sheet in the geologically short time of 10,000 yr. Conversely, high snowfalls will not be preserved if it gets too hot in the summer.

Because the seasonal cycle is so strongly controlled by the land-sea distribution, it is important to include this distribution in palaeoclimate simulations. Previous experiments with three-dimensional general circulation models (GCMs) have not fully explored the role of the seasonal cycle, because about 15 model years are needed to approach equilibrium with a passive mixed-layer ocean. Such a simulation could require as much as 100 hours on a Cray-class supercomputer. Use of the simpler EBM allows us to examine the effect of different boundary conditions with a flexibility not attainable by GCMs.

† Present address: Applied Research Corporation, ARC Technologies, 305 Arguello Drive, College Station, Texas 77840, USA.

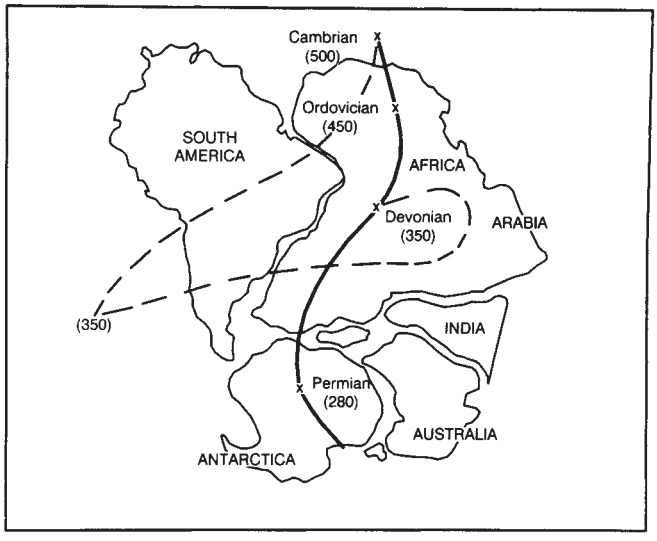


Fig. 1 Postulated Palaeozoic apparent polar wander paths of the South Pole with respect to Gondwanaland (after ref. 25). Dashed line indicates alternate path based on different data sources[25,26]. Numbers in parentheses are ages in Myr

The model

The model is a linear version of a two-dimensional seasonal energy balance model (EBM) described in detail elsewhere[10]. The second dimension enables relatively realistic resolution of land-sea distribution. The model may be summarized by its single governing equation:

$$C(r)\frac{\delta T(r,t)}{\delta t} - \nabla[D(x)\cdot\nabla T(r,t)] + A + BT(r,t) = QS(x,t)a(x)$$

where $T(r, t)$ is the surface temperature field; r is a point on the Earth's surface; x is the sine of the latitude; t is the time of year; $C(r)$ is the effective heat capacity of the Earth-atmosphere column; $D(x)$ is the isotropic heat diffusion coefficient, with smooth latitude dependence; A and B are empirical coefficients representing infrared radiation back to space; Q is the solar constant divided by 4; $S(x, t)$ is the seasonal distribution of solar radiation; and $a(x)$ is the co-albedo of the Earth-atmosphere system, assumed to depend only on latitude.

The model calculates a balance between absorbed solar radiation, diffusive heat transport and outgoing infrared radiation. The temperature response is dependent on a spatially varying heat capacity (low for land, high for water). The geographic resolution of input data is ~5°×5°, comparable to that of more complex GCMs, but the solution method effectively smooths patterns on scales less than ~1,600 km. The model calculates a single temperature for the entire column of the Earth/atmosphere or mixed-layer ocean/atmosphere grid. There is generally very good agreement between model simulations of the present annual cycle and observations[10].

Although many aspects of model sensitivity and parameterizations have been discussed before[11], it is useful to describe some features and assumptions in greater detail in order to assess the level of confidence in the very large temperature changes discussed below. The small number of adjustable parameters in the model are generally well constrained by observation or theory. The albedo term accounts for the latitudinally varying solar zenith angle effect[12], and is based on satellite measurement[13] of the Earth's radiation budget. For mid- and high-latitude land regions, the infrared parameterization[14] is also supported by satellite measurements[15]. For the largest spatial and temporal scales, the diffusion approximation is supported by analysis of atmospheric temperature data[16]. The 60:1 ratio of heat capacity between ocean and land reflects the ratio of the e-folding time for a radiative perturbation of a mixed-layer ocean to that for land (five years, as against one month). The diffusion coefficients were fitted to conform to the present seasonal temperature profile over North America. The latitude dependence of the diffusion coefficients is based on the argument that the tropical Hadley cells are more efficient at smoothing temperature anomalies than are the mid-latitude eddies[17]. Although there have probably been changes in total heat transport and in the partitioning of transport between the oceans and atmosphere, the gradients of total heat transport have remained about the same. This postulate is supported by several GCM experiments[8,18-20] and by theoretical considerations[21,22]. As the model does not distinguish between the temperature of the atmosphere and mixed-layer ocean, past changes in partitioning of transport between the two fluids should not affect the calculations.

Although many feedbacks are not included in the model, the results are considered to be fairly robust. This conclusion is based on several lines of reasoning: (1) the strong correlation between the present seasonal cycle and land-sea distribution[6] indicates a limit to the negative feedback effects of processes not included in our model; (2) comparison of EBM and GCM results indicates comparable sensitivities for summer temperature perturbations in continental interiors[7]; and (3) on a seasonal timescale, temperatures in continental interiors are primarily related to solar forcing[23], and the relation between forcing and response appears to be nearly linear[8,11].

Thus, despite the relative simplicity of the model, we believe that the results provide considerable insight into past changes in the seasonal cycle. We have the simplest form of a realistic model that can provide such information. Given that the land-sea distribution is the best known geological boundary condition, we have a set of experiments which should be of considerable interest. Even though EBMs are relatively low in the hierarchy of climate models[24], the present results may be relatively high in the 'hierarchy of believability' of climate model results.

The experiments discussed below involved no adjustments of parameters other than those involved in tuning the model to the present climate. We have also not allowed for snow-albedo feedbacks, because we first want to understand the effect of seasonal cycle changes without the complication of a large and uncertain feedback. Owing to the limitations of the approach, our experiments can be considered only as sensitivity tests, rather than simulations of past climates.

Idealized geography

Palaeomagnetic data[25] for Gondwanaland (Fig. 1) indicate that the South Pole was offshore during the Cambrian, in North Africa during the Ordovician, and in the vicinity of Antarctica in the late Palaeozoic. The detailed history of the migration from North Africa to Antarctica is a subject of considerable debate. As the primary purpose of this paper is the exploration of mechanisms rather than realistic simulations for different possible configurations, we concentrate on the extreme cases when the pole is located at the edge and centre of the supercontinent.

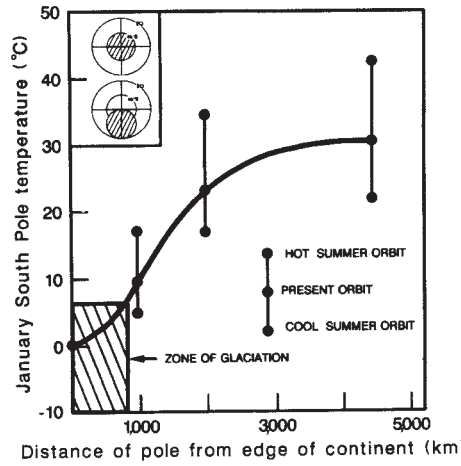


Fig. 2 Summer temperature at the South Pole for different positions of an idealized supercontinent (see inset). Heavy line represents temperatures for the present orbital configuration. Variations around this value represent changes in forcing compatible with the maximum changes in orbital configurations for the Pleistocene[26] (highest points, 'hot summer' orbit; lowest points, 'cool summer' orbit). Diagonal striping marks the postulated region of parameter space affected by glaciation if ice-albedo feedback were included (see text).

To clarify concepts, we first conducted a series of experiments with simplified geography. The area of Gondwanaland is approximately equivalent to that of a circular continent centred on the South Pole and extending to 45° S. Our first set of experiments involved a circular ('disc') continent, with the South Pole initially at the centre, then moved to the edge through two intermediate positions. Three simulations were run for each position, one using the present seasonal insolation cycle, and two using extreme orbital configurations for the Pleistocene[26], to determine the potential effect on climate of Milankovitch insolation perturbations. GCM experiments indicate significant sensitivity of temperature response to variations in orbital insolation[27,28].

When the disc is centred on the South Pole, summer temperatures are ~30 °C (Fig. 2). The magnitude of the warming is so large that it offers support for the hypothesis that an ice-free Earth could be maintained by warm summers alone[7,29]. Such a climate change would be associated with relatively minor changes in the globally averaged annual temperature: an ice-free Earth does not necessarily have to be a particuarly warm Earth. Before the advent of land plants, seasonal temperature increases in continental interiors may have been even greater than modelled here as a result of decreased levels of soil moisture, causing a substantial positive temperature feedback[30,31]. These results imply that the standard model linking ice cap growth to continental positions in high latitudes (see for example, refs 32, 33) may require some modification[34].

When the South Pole is at the edge of the continent, summer temperatures are near freezing and are consistent with a glaciated state. The low temperatures in coastal regions reflect the very large heat capacity of the mixed-layer ocean, which tends to suppress the magnitude of summer warming in nearby land areas. Increased moisture availability in coastal areas would also favour glacier growth. Milankovitch insolation fluctuations produce a range of ~ ±6-10 °C around the mean summer temperature in continental interiors, but only ~2-3 °C in coastal regions. The magnitude of this range in interior regions is comparable to that derived from GCM simulations for the Cretaceous[35], again supporting our contention that the model sensitivity may be comparable to that of GCMs.

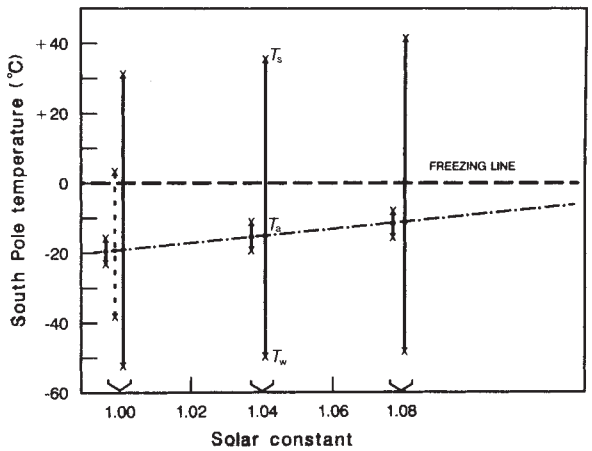


Fig. 3 Effect of heat capacity on the annual range of temperatures at the pole. Dot-dash line, annual-average temperature (T_a) plotted against the value of the solar constant. Summer and winter temperatures (T_s, T_w) are shown for three values of the heat capacity at the poles: high heat capacity (smallest temperature range), as for a marine setting; medium heat capacity (dashes vertical line), for a region covered by permanent sea ice; and low heat capacity (largest temperature range), for a supercontinent centred on the pole. Values for intermediate heat capacity are based on present-day observations for the Arctic Ocean[36].

The effect of heat capacity on the annual range of temperatures at the pole (Fig. 3) indicates that low heat capacity (large land area) can produce an annual range of >80 °C, while high heat capacity (water) potentially restricts the annual range to ~6–8 °C. The annual mean temperature is the same for both cases: −18 °C at the poles for present forcing. In reality, summer temperatures will not decrease to the theoretical limit, because sea water freezes at ~ −2 °C. Sea ice has a heat capacity intermediate between land and water[10], so the annual range of temperatures for a polar oceanic region with sea-ice cover is greater than for open water[36]. Nevertheless, summer temperatures are still low, and the high heat capacity of water would keep them low even if the sea ice were to melt. It is evident from Fig. 3 that, in order to have summer temperatures at the pole significantly above zero for a marine setting, an annually averaged temperature of ~0 °C is required. This value is ~20 °C greater than the present annual mean. Our calculations therefore predict sea ice for open-water sites in polar regions (see also ref. 37). The low temperatures in coastal regions would also favour ice formation on nearby land areas.

Previous experiments with a nonlinear version of the model[38] suggest that when summer temperatures decrease to 4–6 °C above freezing, ice-albedo feedback effects will force the climate into a glaciated state, perhaps as a result of an ice-cap instability mechanism[38–40]. The diagonally striped region of parameter space in Fig. 2, at $T<6$ °C, is therefore considered to be the 'zone of glaciation'. Based on these calculations, we hypothesize that when the South Pole is within ~800 km of the edge of the supercontinent, glaciation is possible. When it is slightly farther away, glaciation is most probable during Milankovitch 'cool summer' orbits (lowest points in Fig. 2). When it is much more than 1,000 km distant, summer temperatures may generally be too warm to allow glaciation.

Increased solar constant

In a further set of experiments, the solar constant was increased by 4% and 8% in order to mimic a CO_2 increase. The regional climate response to solar constant and CO_2 fluctuations have been shown to be comparable in GCM experiments[41,42]. Although a 2% change in the solar constant is approximately equivalent to a doubling of CO_2, 4% and 8% increases in the solar constant are equivalent to five-fold and thirty-fold increases in CO_2 radiative forcing. The logarithmic relationshp[43] reflects the saturation of different CO_2 absorption bands at higher concentrations. A four- to ten-fold increase in CO_2 has sometimes been cited as an estimated level required for ice-free states[4,44].

Although the linear model compares favourably to GCMs for seasonal insolation changes over mid-latitude land masses, the comparison may not be so good for changes in the annual-average radiation budget. Atmospheric feedback effects in GCMs (such as those resulting from water vapour and clouds) effectively double the sensitivity of the global temperature response to variations in the solar constant or CO_2[45,46]. In order to account crudely for additional feedbacks in net annual changes in the radiation budget, we used the very large change in solar constant forcing of 8%. Despite the limitations of the model, we believe our simulations have some merit, owing to the linkage between the large heat capacity of the mixed-layer ocean and temperatures in nearby coastal areas.

In general, increasing the solar constant has a greater effect on the mean annual temperature than on the seasonal cycle (Fig. 3). Because of the high heat capacity of water, it is necessary to raise the mean annual temperature to about zero in order to obtain ice-free solutions for a polar region with open water. A very large change in forcing is required to achieve this effect. In all likelihood, the required change in forcing is less than calculated here, because we do not consider feedbacks for annually averaged forcing. But the important point is that it may be possible to have a glaciated state with significantly higher CO_2 levels, if the pole is located in coastal regions.

The above 'CO_2' experiment may resolve a paradox pointed out by Van Houten[47] for the late Ordovician (~450 Myr), namely that the abudant evidence for glaciation[48] is seemingly incompatible with geological evidence suggestive of higher CO_2 levels[49–53], with values perhaps as high as those of the mid-Cretaceous. Our study suggests that it may be possible to resolve the high CO_2/glaciated Earth paradox if the seasonal cycle is taken into account. The paradox is resolved in a manner entirely consistent with present climate theory and requires no elaborate modification of climate models.

Realistic geography

Simulations with realistic geographies substantiate the basic conclusions derived from the idealized geography experiments, and provide additional insight into regional variations in temperature. We chose two time intervals that capture the potential extremes of climate and continentality for the Palaeozoic—the early Devonian (Emsian, ~390 Myr) and the late Ordovician (Caradocian, ~450 Myr). Unfortunately, the precise position of the early Devonian South Pole is open to question[25,54] (Fig. 1). As we are primarily concerned with illustrating a concept, we concentrate on the two configurations representative of the potential extremes for the idealized geography case, recognizing that a full range of experiments is required to explore the alternative plate reconstructions. We used the base maps of ref. 55 and superimposed the sea-level records of ref. 56 on all reconstructions.

For the realistic geography experiments (Fig. 4), flooding of continental margins by shallow seas effectively reduces the size of the supercontinent by almost 30%. Simulated January tempeatures at the South Pole were 21 °C for a centrally located pole location[57], and ~0 °C for the more coastally located pole location. These simulations are consistent with geological data indicating no ice for a 60-Myr time interval centred on the early Devonian, and for glaciation in North Africa during the late Ordovician[48], The early Devonian results imply that it is not necessary to have a South Pacific pole position (see also Fig. 1) for this time interval in order to explain the ice-free state[48].

Although temperatures at the South Pole are high for the early Devonian plate configuration, they are lower in coastal regions. This pattern is very similar to the one illustrated in ref. 58 for the late Palaeozoic, in which the locus of ice sheets was

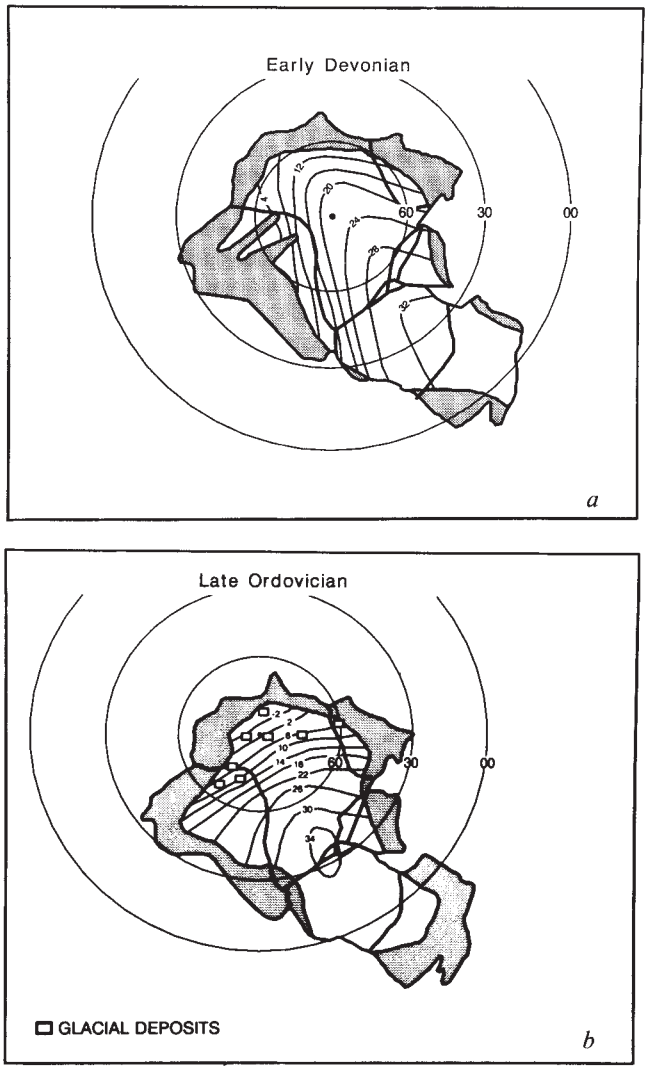


Fig. 4 Realistic geography simulations for the early Devonian (*a*) and late Ordovician (*b*). Circular contours, lines of latitude; other contours, January temperatures. Shaded areas represent continental margins flooded by shallow seas. Rectangles in *b* are glacial deposits from ref. 48. Early Devonian pole position from ref. 57 and land-sea distributions from ref. 56. Late Ordovician pole position from ref. 55. In *b*, the land-sea distribution was adapted from a middle Ordovician reconstruction in ref. 56 and assumes that the area affected by shallow seas did not change significantly between the middle and late Ordovician.

shifted away from the pole to more coastal locations. The general patterns described above also apply to an alternative early Devonian reconstruction (not shown) with a southern African pole[59], except that January temperatures at the South Pole are slightly cooler (18 °C).

The realistic-geography experiments also provide information about climate in low latitudes. The extension of the supercontinent into low lattitudes opens up a corridor of warmth into high-latitude sectors that are contiguous with the low-latitude extension. January temperatures are quite high in low latitudes for the Ordovician simulation, with the maximum temperature of 34 °C greater by ~2 °C than the summer maximum simulated by the model for the present climate[10]. This result is suggestive of generally arid conditions in continental interiors, with perhaps a strong monsoon in coastal regions[27,28].

Late Palaeozoic glaciations

The lessons learned from these simulations can be applied to the late Palaeozoic glaciations in 'southern' Gondwanaland. The onset of widespread glaciation[58] in Africa, South America, Antarctica and India can be interpreted in terms of a coastal (Antarctic) location for the South Pole (Fig. 1). Decreasing CO_2 levels[60] could also have contributed to glacial expansion. But as our analysis suggests, the land-sea distribution may be more important than CO_2 in dictating the time of glacial onset. Our simulations also allow us to predict favourable sites for glacial initiation. The general tendency for isotherms in high latitudes to be aligned parallel to coastlines (see Fig. 4) would suggest that glaciation could occur anywhere along a strip extending from South America to Australia. Mountain-building events[61] might favour some areas over others in this strip (our model does not resolve topography). Finally, as the pole moved offshore in the late Permian, continental glaciation would cease. However, there may still have been sea ice in polar regions[37] (see Fig. 3).

Implications for origin of ice-free climates

Our two most important results are that summer temperatures in the interior of large polar land masses may sometimes be high enough to prevent ice-cap formation, and that it may be possible to have a glaciated Earth, even with higher CO_2 levels, if the pole is located in coastal regions. As we discuss in the model section, several lines of evidence support the basic robustness of these model results, as do preliminary experiments with a nonlinear version of our model. Still, it would clearly be desirable to test our conclusions with more physically comprehensive models. We note, however, that even GCM simulations may be associated with large uncertainties[46], especially where sea-ice formation is involved. Nevertheless, the magnitude of the simulated temperature changes are sufficiently large to support our proposition that satisfactory simulations of the history of Phanerozoic temperatures should include the seasonal cycle.

Given that it may be possible to have an ice-free Earth without changing CO_2 levels, or a glaciated Earth with significantly higher CO_2 levels, our simulations raise questions about the presumed pre-eminent role of CO_2 in long-term climate change. Furthermore in some cases neither CO_2 nor seasonality can explain warmer temperatures in high latitudes[7,8]. Changes in ocean heat transport may also have to be considered[8,62-64]. This analysis implies that the long-term evolution of climate involves processes more complicated than generally assumed. Descriptions of the average climate state of an interval should distinguish between seasonal and geographic changes in temperatures, and year-round warmth. This conclusion also has implications for geochemical models, which sometimes parameterize relationships between weathering rates and mean climate state.

Supported by funds from the NASA Climate Research Program and an NRC fellowship to T.J.C. We thank R. A. Berner, J. C. Crowell, D. V. Kent, G. R. North, D. K. Rea and M. E. Suarez for comments on an earlier draft, D. V. Kent and R. B. Hargraves for discussions on Palaeozoic palaeomagnetism, and R. Evans for assistance in drafting.

Received 20 April; accepted 14 September 1987.

1. Barron, E. J., Thompson, S. L. & Schneider, S. H. *Science* **212**, 501-508 (1981).
2. Crowley, T. J. *Rev. Geophys. Space Phys.* **21**, 828-877 (1983).
3. Barron, E. J. & Washington, W. M. *J. geophys. Res.* **89**, 1267-1279 (1984).
4. Barron, E. J. & Washington, W. M. in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 546-553 (American Geophysical Union, Washington, 1985).
5. Berner, R. A., Lasaga, A. C. & Garrels, R. M. *Am. J. Sci.* **283**, 641-683 (1983).
6. Wallace, J. M. & Hobbs, P. V. *Atmospheric Science: An Introductory Survey* (Academic, New York, 1977).
7. Crowley, T. J., Short, D. A., Mengel, J. G. & North, G. R. *Science* **231**, 579-584 (1986).
8. Crowley, T. J. in *Physically-Based Modelling and Simulation of Climate and Climatic Change* (ed. Schlesinger, M. E.) (Reidel, Dordrecht, in the press).
9. Crowell, J. C. *Am. J. Sci.* **278**, 1345-1372 (1978).
10. North, G. R., Mengel, J. G. & Short, D. A. *J. geophys. Res.* **88**, 6576-6586 (1983).
11. North, G. R., Cahalan, R. F. & Coakley, J. A. *Rev. Geophys. Space Phys.* **19**, 91-121 (1981).
12. Lian, M. S. & Cess, R. D. *J. atmos. Sci.* **34**, 1058-1062 (1977).
13. Stephens, G. L., Campbell, G. G. & Vonder Haar, T. H. *J. geophys. Res.* **86**, 9739-9760 (1981).
14. Budyko, M. L. *Tellus* **21**, 611-619 (1969).
15. Short, D. A., North, G. R., Bess, T. D. & Smith, G. L. *J. Clim. Appl. Met.* **23**, 1222-1233 (1984).

16. Lorenz, E. N. *J. atmos. Sci.* **36**, 1367-1376 (1979).
17. Lindzen, R. S. & Farrell, B. *J. atmos. Sci.* **34**, 1487-1501 (1977).
18. Barron, E. J. & Washington, W. M. *Geology* **10**, 633-636 (1982).
19. Manabe, S. & Bryan, K. *J. geophys. Res.* **90**, 11689-11708 (1985).
20. Rind, D. *J. atmos. Sci.* **43**, 3-24 (1986).
21. Stone, P. H. *Dyn. Atmos. Oceans* **2**, 123-139 (1978).
22. Covey, C. & Barron, E. J. *Earth Sci. Rev.* (in the press).
23. Otto-Bliesner, B. L. & Houghton, D. D. *J. geophys. Res.* **91**, 6682-6694 (1986).
24. Schneider, S. H. & Dickinson, R. *Rev. Geophys. Space Phys.* **12**, 447-493 (1974).
25. Morel, P. & Irving, E. *J. Geol.* **86**, 535-561 (1978).
26. Berger, A. L. *Quat. Res.* **9**, 139-167 (1978).
27. Kutzbach, J. E. & Street-Perrott, F. A. *Nature* **317**, 130-134 (1985).
28. Kutzbach, J. E. & Guetter, P. J. *J. atmos. Sci.* **43**, 1726-1759 (1986).
29. Watts, R. G. & Hayder, M. *Ann. Glaciol.* **5**, 234-236 (1984).
30. Shukla, J. & Mintz, Y. *Science* **215**, 1498-1501 (1981).
31. Manabe, S. & Wetherald, R. T. *Science* **232**, 626-628 (1986).
32. Crowell, J. C. & Frakes, L. C. *Am. J. Sci.* **268**, 193-224 (1970).
33. Frakes, L. A. *Climates Throughout Geologic Time* (Elsevier, Amsterdam, 1979).
34. Burrett, C. F. *Nature* **296**, 54-56 (1982).
35. Glancy, T. J., Barron, E. J. & Arthur, M. A. *Paleoceanography* **1**, 523-538 (1986).
36. Crutcher, H. L. & Meserve, J. M. *Selected Level Heights, Temperatures and Dew Points for the Northern Hemisphere* (Naval Weather Service Command, Washington)
37. Hunt, B. G. *Nature* **308**, 48-51 (1984).
38. North, G. R. & Crowley, T. J. *J. geol. Soc., Lond.* **142**, 475-482 (1985).
39. Suarez, M. J. & Held, I. M. *J. geophys. Res.* **84**, 4825-4836 (1979).
40. North, G. R. *J. atmos. Sci.* **41**, 3390-3395 (1984).
41. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **37**, 99-118 (1980).
42. Hansen, J. *et al.* in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 130-163 (American Geophysical Union, Washington, 1984).
43. Madden, R. A. & Ramanathan, V. *Science* **209**, 763-768 (1980).
44. Lasaga, A. C., Berner, R. A. & Garrels, R. M. in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 397-411 (American Geophysical Union, Washington, 1985).
45. North, G. R., Mengel, J. G. & Short, D. A. in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 164-170 (American Geophysical Union, Washington, 1984).
46. Schlesinger, M. E. & Mitchell, J. F. B. in *Projecting the Climatic Effects of Increasing Carbon Dioxide* (eds MacCracken, M. C. & Luther, F. M) 81-147 (US Dept Energy, Washington, 1985).
47. Van Houten, F. B. *Geology* **13**, 722-724 (1985).
48. Caputo, M. V. & Crowell, J. C. *Geol. Soc. Am. Bull.* **96**, 1020-1036 (1985).
49. Vail, P. R., Mitchum, R. M. & Thompson, S. *Mem. Am. Ass. Petrol. Geol.* **26**, 83-97 (1977).
50. Hallam, A. *A. Rev. Earth planet. Sci.* **12**, 205-243 (1984).
51. Sanders, P. A. *Nature* **305**, 19-22 (1983).
52. Wilkinson, B. H. & Given, R. K. *J. Geol.* **94**, 321-333 (1986).
53. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton University Press, 1984).
54. Kent, D. V. & May, S. R. *Rev. Geophys.* (in the press).
55. Smith, A. G., Hurley, A. M. & Briden, J. C. *Phanerozoic Paleocontinental World Maps* (Cambridge University Press, 1981).
56. Ziegler, A. M. *et al. A. Rev. Earth planet. Sci.* **7**, 473-502 (1979).
57. Hurley, N. F. & Van der Voo, R. *Geol. Soc. Am. Bull.* **98**, 138-146 (1987).
58. Bambach, R. K., Scotese, C. R. & Ziegler, A. M. *Am. Scient.* **68**, 26-38 (1980).
59. Ziegler, A. M. *et al.* in *Paleobotany, Paleoecology, and Evolution* (ed. Niklas, K. J.) 231-266 (Praeger, New York, 1981).
60. Fischer, A. G. in *Climate in Earth History* (ed. Berger, W.) 97-104 (National Academy Press, Washington, 1982).
61. Powell, C. McA. & Veevers, J. *Nature* **326**, 177-179 (1987).
62. Savin, S. M., Douglas, R. G. & Stehli, F. G. *Geol. Soc. Am. Bull.* **86**, 1499-1510 (1975).
63. Kennett, J. P. *J. geophys. Res.* **82**, 3843-3860 (1977).
64. Shackleton, N. J. & Boersma, A. *J. Geol. Soc., Lond.* **138**, 153-157 (1981).
65. Schmidt, P. W. *et al. Tectonics* **5**, 135-150 (1986).

LETTERS TO NATURE

Cosmic rays and gamma radiation from the shell of SN1987A

V. S. Berezinsky* & V. L. Ginzburg†

*Institute for Nuclear Research of the Academy of Sciences of the USSR, 60th Anniversary of October Revolution prospect 7*a*, Moscow 117312, USSR
†P. N. Lebedev Physical Institute of the Academy of Sciences of the USSR, Leninsky prospect 53, Moscow 117924, USSR

Detection of high-energy gamma rays from young supernovae shells[1,2] can directly prove the hypothesis that the main sources of cosmic rays (CR) in our Galaxy are supernovae. This radiation is produced in nuclear collisions of accelerated protons and nuclei, through the decay of pions. On 13 April 1987 an attempt was made to measure the gamma radiation from SN1987A between 50 and 500 MeV in energy by an international team from Australia, UK, FRG and USA (R. Stauberg, personal communication). Spark chamber measurements from a balloon gave a preliminary upper limit to the flux of $j_\gamma < 3\times10^{-4}\,\mathrm{cm^{-2}\,s^{-1}}$. The search for high-energy gamma rays is also possible using the ground-based Cerenkov-light detectors at Potchefstroom (S. Africa) and White Cliff station (Australia) for $E_\gamma \gtrsim 1$ TeV, and by means of the extensive air shower (EAS) array at Buckland Park (Australia) for $E_\gamma \gtrsim 10^6$ GeV. Such observations, we show here, can discover CR in the SN1987A shell if they are produced inside the shell with luminosity down to $L_p \sim 10^{39}\,\mathrm{erg\,s^{-1}}$. This can support or reject a very wide class of the models of CR production by supernovae. We argue that such measurements for SN1987A will be possible during the next 1–2 years, enough time to move Cerenkov detectors from the Northern to the Southern Hemisphere.

The total energy output W_p^{SN} per supernova in protons with kinetic energy $E_p \gtrsim 1$ GeV can be found phenomenologically[3] from the observed CR energy density, $W_p(E_p > 1\,\mathrm{GeV}) \approx 1.4\times10^{-12}\,\mathrm{erg\,cm^{-3}}$, the mass of gas in the Galaxy, $M_g \approx 6\times10^9 M_\odot$, the mean pathlength of CR, $\bar{x} \approx 5\,\mathrm{g\,cm^{-2}}$, and the SN explosion frequency in Galaxy, $\nu_{SN} \approx 0.03\,\mathrm{yr^{-1}}$:

$$W_p^{SN} \approx \frac{cW_pM_g}{\nu_{SN}\bar{x}} = 10^{50}(M_g/6\times10^9 M_\odot)(0.03/\nu_{SN}) \qquad (1)$$

From observations of the spectrum, anisotropy and secondary-to-primary nuclei ratio, the CR injected into the Galaxy must be described[3] by a power law spectrum with index $\gamma = 1.1-1.4$ up to energy $E \sim 10^6$ GeV, perhaps up to $E \sim 10^{10}$ GeV. Then the number of protons with kinetic energy E produced by one supernova is

$$N_p^{SN}(E)\,dE = (\gamma-1)(W^{SN}/E_0^2)(E/E_0)^{-(\gamma+1)} \qquad (2)$$

For simplicity we assume that $N_p^{SN}(E) = 0$ at $E < E_0$, thus implying that W_p^{SN} comes from protons with $E \geqslant E_0$, and that $E_0 \approx 0.5$-1 GeV.

While W_p^{SN} is empirically determined rather reliably by equation (1), the CR luminosity, $L_p^{SN} = W_p^{SN}/\tau_{acc}$, depends on the duration of the acceleration phase, τ_{acc}, of supernova and thus is model dependent.

If CR are accelerated at the outer shock front, the acceleration phase is long, with $\tau_{acc} \gtrsim 10^3$-10^4 yr. Moreover, the accelerated particles are then confined between the outer and reflected shock fronts and never penetrate the dense part of the shell. In this model the expected gamma-radiation is weak, and it will not be discussed here.

Another class of model is connected with acceleration inside the shell. The most probable mechanisms are acceleration by a pulsar and Fermi acceleration by magnetic turbulence. In this case the intrinsic CR luminosity (that of the 'accelerator') is $L_p^{SN} \approx AW_p^{SN}/\tau_{acc}$, where A describes the adiabatic energy losses in the expanded shell. This factor can be ~1 in some exotic models where there is compensation of adiabatic losses and acceleration[4], but can be up to $\sim10^3$ in general. The acceleration phase must be rather short. Its value can be derived from the shape of the supernova light curve. Typically, at $t_0 \approx 100$ d after explosion the optical luminosity decreases exponentially as $L(t) = L_0 \exp(-(t-t_0)/\tau)$. The CR luminosity decreases much more slowly. In the late stages of expansion, when the shell is transparent to gamma radiation, neutrinos and gamma rays, produced via proton-nucleon (pN) collisions and pion decays, freely leave the shell, while a large part of the energy of secondary electrons and positrons is thermalized and radiates as black-body radiation from the surface of the photosphere. Taking into account the fact that the probability of a nuclear collision diminishes as t_a^2/t^2 with time (see below) and assuming that CR luminosity L_p is constant during τ_{acc}, it is possible to derive τ_{acc} from the condition that CR must not produce a noticeable flat step in the light curve[5]:

$$0.05(\gamma+1)^{-1}(1-\alpha^\gamma)^{-1}(W_p^{SN}t_a^2/\tau_{acc}^3) \leqslant L_0 \exp(-(\tau_{acc}-t_0)/\tau) \qquad (3)$$

where $\alpha \approx 0.5$ is the energy retained by a nucleon in nuclear collision. For $t_0 \approx \tau \approx 100$ d and $L_0 \approx 1 \times 10^{42}$ erg s^{-1}, as in SN1987A, $\tau_{acc} < 1$–10 yr, giving $L_p^{SN} > 10^{41}$ erg s^{-1}.

As an upper limit we take $L_p^{SN} \lesssim 10^{43}$ erg s^{-1} which applies in models[6] where the shell is ejected by pulsar radiation. Thus, we conclude that

$$10^{41}\ \text{erg s}^{-1} \lesssim L_p^{SN} \lesssim 10^{43}\ \text{erg s}^{-1} \tag{4}$$

The generation of gamma-photons, neutrinos, neutrons, positrons and antiprotons has been considered in refs 1, 2, 7–14. From the above, the number $Q(E, t)$ of accelerated protons with kinetic energy E injected per unit time in the expanded shell is given by

$$Q(E, t) = (\gamma - 1)(L_p(t)/E_0^2)\varepsilon^{-(\gamma+1)} \tag{5}$$

(and $Q(E, t) = 0$ for $\varepsilon < 1$) where $\varepsilon = E/E_0$ is dimensionless energy and $L_p(t) = \int EQ(E, t)\,dE$ is CR luminosity. The kinetic equation for the total number, $N_p(\varepsilon, t)$, of protons with energy ε in the shell is

$$\begin{aligned}\frac{\partial N_p(\varepsilon, t)}{\partial t} &= Q(\varepsilon, t) - c\sigma_{pN}\bar{n}_N(t)N_p(\varepsilon, t) \\ &+ c\sigma_{pN}\bar{n}_N(t)\int_\varepsilon^\infty N_p(\varepsilon' t) \\ &\frac{1}{\sigma_{pN}}\frac{d\sigma_{pN}(\varepsilon', \varepsilon)}{d\varepsilon}d\varepsilon' \\ &+ \frac{\partial}{\partial\varepsilon}(\beta_a(\varepsilon)N_p(\varepsilon, t)) = 0\end{aligned} \tag{6}$$

where the second and third terms describe nuclear collisions of the protons (σ_{pN} is pN cross-section) and the fourth term accounts for adiabatic energy losses $d\varepsilon/dt = -\beta_a(\varepsilon)$.

We take the energy distribution of recoil nucleons in pN-scattering to be

$$\sigma_{pN}^{-1}\,d\sigma_{pN}(\varepsilon', \varepsilon)/d\varepsilon = \delta(\varepsilon - \alpha\varepsilon') \tag{7}$$

where α is the mean fraction of energy retained by the incident proton in nuclear collision ($\alpha \approx 0.5$). Use of a distribution $\sigma_{pN}^{-1}\,d\sigma_{pN}(\varepsilon', \varepsilon)/d\varepsilon = 1/\varepsilon'$, instead of equation (7), results in a 2–7% difference in the solution of equation (7) for $1.1 \lesssim \gamma \lesssim 1.6$.

The accelerated protons and the gas are not uniformly mixed in the shell. A Rayleigh–Taylor instability develops as a cloud-like structure of the shell and produces low density regions, bubbles filled by the 'light gas', that is, by cosmic rays. Moreover, CR penetrate through the low density gaps to the periphery of the shell, where the gas density is small. CR carry with them the frozen magnetic field and the ionized gas and this conglomerate pushes out the intersellar gas. Equation (7) contains $\bar{n}_N(t)$, the average number density of gas in the bubbles:

$$\bar{n}_N(t) = \eta \cdot 3M/(4\pi m_N \bar{u}^3 t^3) \tag{8}$$

where η denotes the mixing of CR and gas ($\eta < 1$; probably $\eta \ll 1$), M is the mass of the shell and $\bar{u} \equiv \sqrt{\langle u^2 \rangle}$ is the r.m.s. velocity of the shell. To calculate $\langle u^2 \rangle$ we used the model of Grassberg *et al.*[15] which successfully describes the light curve of SN1987A if $M = 16M_\odot$. The density distribution $\rho(u)$ in the shell was taken from ref. 16. This gave $\bar{u} \approx 5 \times 10^8$ cm s^{-1}.

In free expansion during which the mass M_{cap} of the captured interstellar gas remains less than the ejected mass M_{ej}, the velocity $\bar{u}$ is constant and adiabatic energy loss is $\beta_a(\varepsilon) = \varepsilon/t$. With all these assumptions equation (7) can be solved for $\varepsilon > 1$, giving

$$N_p(\varepsilon, t) = (\gamma - 1)(W_p(t)/E_0^2)\varepsilon^{-(\gamma+1)} \tag{9}$$

where $W_p(t)$ is the total energy in the shell of protons whose individual energies are higher than E_0. (Due to energy losses there are now protons in the shell with $E < E_0$.) From equation (7) it is easy to see that

$$dW_p(t)/dt = L_p(t) - (t_a^2/t^3)W_p(t) - (\gamma/t)W_p(t) \tag{10}$$

where

$$\begin{aligned}t_a &= \left[\frac{3\eta Mc\sigma_{pN}}{4\pi m_N \bar{u}^3}(1 - \alpha^\gamma)\right]^{1/2} \\ &= 1.5 \times 10^8 \eta^{1/2}\left(\frac{M}{20M_\odot}\right)^{1/2}\left(\frac{5 \times 10^8}{\bar{u}}\right)^{3/2}\end{aligned} \tag{11}$$

defines the duration of nuclear phase (when nuclear energy losses dominate over adiabatic energy losses). The protons injected in the shell at $t < t_a$ lose all their energy in nuclear collisions.

The solution of equation (7) during the free expansion ($M_{cap} < M_{ej}$) can now be easily found for the slowly declining CR luminosity $L_p(t) \approx L_p$ as

$$W_p(t) = L_p t_a f(t), \tag{12}$$

$$f(t) = (t/t_a)^{-\gamma}\exp(t_a^2/2t^2)\int_0^{t/t_a} dx \cdot x^\gamma \exp(-x^{-2}/2)$$

when $t \ll t_a$, $f(t) \approx (t/t_a)^3$, and when $t \gg t_a$, $f(t) \approx t/(\gamma + 1)t_a$.

To calculate the total gamma ray flux from π^0 decays following pN-collisions, producing photons mainly with energies 70–1,000 MeV, we begin at time t_γ when the shell becomes transparent to high energy gamma rays. If this happens when the mean thickness of the shell becomes equal to the radiation length x_{rad} (≈ 70 g cm^{-2} for a shell with relative hydrogen abundance $X = 0.73$, $Y = 0.25$ for helium), then

$$t_\gamma = \left(\frac{3M\eta_1}{4\pi\bar{u}^2 x_{rad}}\right)^{1/2} = 2.3 \times 10^7 \eta_1^{1/2}\left(\frac{M}{20M_\odot}\right)^{1/2}\left(\frac{5 \times 10^8}{\bar{u}}\right) \tag{13}$$

where $(1 > \eta_1 > \eta)$ because the CR bubbles are separated by dense gas clouds.

One might think that due to electromagnetic cascading the photons can traverse a thickness $x > x_{rad}$ and therefore that t_γ can be smaller, but it can be shown[5] that for the total number of photons from π^0 mesons this is not the case, especially if the spiralling of electrons in the magnetic field is taken into account. Inclusion of the electromagnetic cascade does not increase the total number of photons at $x > x_{rad}$.

The total gamma-ray flux from π^0 decays per unit time at the nuclear stage of expansion is

$$F_\gamma(t) = 2c\bar{n}_N(t)\int_{E_{th}}^\infty N_p(E, t)\sigma_{\pi^0}(E)\,dE \tag{14}$$

where σ_{π^0} is the inclusive cross-section for π^0 production ($p + N \to \pi^0 + \cdots$) and E_{th} is the threshold energy for pion production. Note that the mixing parameter η is included in $\bar{n}_N(t)$. Inserting equations (9) and (12) into (14) one obtains for $t < t_a$

$$F_\gamma(t) = 2\frac{\sigma_{eff}}{\sigma_{pN}}\frac{1}{1 - \alpha^\gamma}\frac{\gamma - 1}{\gamma}\frac{L_p}{E_0} \tag{15}$$

where $\sigma_{eff} = \gamma\int_{E_{th}/E_0}^\infty d\varepsilon \cdot \varepsilon^{-(\gamma+1)}\sigma_{\pi^0}(\varepsilon)$ is the spectrally averaged cross-section for E_{th}/E_0 π^0 production. Formula (15) has a clear physical meaning: the factor 2 comes from the two photons in π^0 decay, $(\gamma - 1)L_p/\gamma E_0$ is the total number of protons injected per unit time in the shell, σ_{eff}/σ_{pN} is the probability of π^0 production in one nuclear collision and $(1 - \alpha^\gamma)^{-1}$ takes into account repeated collisions of recoil nucleons. In this approximation the mixing factor η does not enter into the formula, because $F_\gamma \approx \bar{n}_N(t)N_p(\varepsilon, t)$, with $\bar{n}_N(t) \approx \eta$ and $N_p(\varepsilon, t) \approx 1/\eta$ at $t < t_a$ (see equation (12)). The latter relation is easy to understand because the equlibrium number of protons in the shell, $N_p(\varepsilon, t)$, is defined by nuclear energy losses, and the smaller these losses (through decreased η), the greater $N_p(\varepsilon, t)$.

For numerical estimates we used $E_0 = 400$ MeV and the approximation for $\sigma_{\pi^0}(\varepsilon)$ given by Dermer[17]. The calculated

Table 1 The flux of gamma rays with $E_\gamma \geqslant 70$ MeV from SN1987A at the nuclear stage of expansion $t_\gamma \leqslant t \leqslant t_a$, for $L_p = 1\times10^{41}$ erg s^{-1}

Spectral index γ	1.1	1.2	1.3	1.4	1.5	1.6
Flux $10^5 j_\gamma$/cm^{-2} s^{-1}	3.0	4.6	5.5	5.9	6.1	6.1

gamma ray flux at Earth is given in Table 1 for the assumed luminosity $L_p = 1\times10^{41}$ erg s^{-1} and a distance to SN1987A of 1.6×10^{23} cm. These fluxes are given for the case of full gamma transparency. The time t_γ when this becomes true is poorly defined numerically by equation (13) because of the mixing factor η_1, and this hampers comparison with the balloon observation of 13 April 1987. If at the time of observation ($t \approx 4.3\times10^6$ s) the shell was indeed transparent to gamma rays (implying $\eta_1 < 0.04$), the observed upper limit $j_\gamma < 3\times10^{-4}$ cm^{-2} s^{-1} and the predicted flux (Table 1) gives an upper limit to L_p for each value of γ; $L_p < 5\times10^{41}$ erg s^{-1} for $\gamma \leqslant 1.3$, for example. The perhaps puzzling increase of flux with γ in Table 1 occurs because the luminosity L_p is fixed in our calculations. At a given L_p, the total number of protons injected into the shell, $(\gamma-1)L_p/\gamma E_0$, increases with γ, giving a greater gamma ray flux.

The flux of gamma radiation with $E_\gamma \geqslant 1$ TeV produced during the nuclear stage $t_\gamma \leqslant t \leqslant t_a$ is[9]

$$F_\gamma(>\varepsilon) = \varphi_\gamma \frac{\gamma-1}{\gamma(1-\alpha^\gamma)} \frac{\sigma_{pN}^{in}}{\sigma_{pN}^{tot}} \frac{L_p}{E_0} \varepsilon^{-\gamma} \qquad (16)$$

where σ_{pN}^{in} and σ_{pN}^{tot} are the inelastic and total pN cross-section and φ_γ the scaling yields, takes the values listed in Table 2. The gamma ray flux at $E_\gamma \geqslant 1$ TeV from SN1987A, assuming $L_p = 10^{39}$ erg s^{-1} and $E_0 = 1$ GeV, is given in the same table. The listed fluxes for small $\gamma \leqslant 1.3$ are detectable by the Cerenkov telescope in Potschefstroom, which can therefore discover CR in the SN1987A shell at luminosity $L_p \approx 10^{39}$–10^{40} erg s^{-1}.

We expect that the gamma transparency stage occurs later than for $E_\gamma \geqslant 70$–100 MeV. For $E_\gamma \geqslant 1$ TeV, parent protons with $E_p \geqslant 10$ TeV are probably not confined in the bubbles and therefore mixing with gas is greater. If protons at these energies are accelerated by a pulsar, the total thickness of the shell is a better measure of the matter traversed by gamma rays than the radiation length used for $E_\gamma \geqslant 70$ MeV. The thickness x (in g cm^{-2}) of the shell of SN1987A is[15,16]

$$x(t) \approx 8.3\times10^2 (100/t)^2 \qquad (17)$$

where t is in days. Nonuniform structure in the gas means that the thickness in some directions is smaller than in others, and therefore in calculating t_γ one must introduce mixing $\eta' < 1$:

$$t_\gamma \approx 280(105/x_{rad} + x_{nuc})^{1/2}(\eta')^{1/2} \text{ days} \qquad (18)$$

where $x_{rad} \approx 70$ g cm^{-2}, and $x_{nuc} \approx 35$ g cm^{-2} is the interaction length of a proton in the shell with $E_p \geqslant 10$ TeV. The electromagnetic cascade plays a negligible role at these energies because, in the presence of a magnetic field with $B \geqslant 0.1$ G, electron energy loss is dominated by synchrotron radiation. (The nuclear range of a primary proton diminishes t_γ to some extent).

Therefore, t_γ for gamma radiation with $E_\gamma \geqslant 1$ TeV is not much smaller than the time, τ_{acc}, when the effective acceleration ceases. It leaves only a short period for the most useful observation of SN1987A. The latest data on the light curve can be used to improve the relation between t_γ and τ_{acc}.

Finally, we shall briefly discuss other kinds of high energy radiation generated in the shell — neutrinos with $E_\nu \geqslant$ 30–100 GeV, gamma rays with $E_\gamma \geqslant 3\times10^6$ GeV and neutrons with $E \geqslant 6\times10^9$ GeV, which can reach Earth without decay. General discussion of these radiations from young supernova shells is in refs 9, 10, 17. The high energy neutrino flux from SN1987A was recently discussed by Gaisser and Stanev[19]. Here we present the results of our calculations[5].

Table 2 The gamma ray yields and fluxes at $E_\gamma \geqslant 1$ TeV from SN1987A, for $L_p = 10^{39}$ erg s^{-1}

Index γ	1.1	1.2	1.3	1.4	1.5	1.6
Yield φ_γ	0.110	0.091	0.071	0.056	0.045	0.036
Flux $10^{11} j_\gamma$/cm^{-2} s^{-1}	1.5	1.1	0.55	0.26	0.12	0.05

High energy neutrinos with $E_\nu \geqslant 30-100$ GeV can be observed using existing underground detectors (areas $\approx$ 100–200 m^2) at the level of 10–30 events per yr. This demands a CR luminosity $L_p \approx 10^{43}$ erg s^{-1} for the case of flat spectra, $\gamma \approx 1.1$–1.3.

Gamma radiation with $E_\gamma \geqslant 3\times10^6$ GeV can be observed by EAS array at Buckland Park[20] in Australia. This array detected the source LMCX−4 at $E_\gamma \geqslant 3\times10^6$ GeV in the Large Magellanic Cloud[21]. Our calculations, with absorption of gamma radiation on microwave background radiation taken into account, have shown that in the case of SN1987A detection is possible ($j_\gamma \approx 3\times10^{-15}$–$5\times10^{-14}$ cm^{-2} s^{-1}) if $L_p \approx 10^{42}$ erg s^{-1} and $\gamma \approx 1.1$–1.3.

The production of ultra high energy ($E > 6\times10^9$ GeV) neutrons in SN1987A was estimated on the basis of the observed diffuse flux of CR at $E \approx 10^{10}$ GeV in the Galaxy. The predicted flux $j_n \approx 2\times10^{-12}$ cm^{-2} for 300 days of observations is too small to be detected by the Buckland Park array.

In summary, if SN1987A has the properties of a typical galactic CR source accelerating a number $F(E) = (\gamma-1)L_p E^{-(\gamma+1)}$ per unit time, the intrinsic CR luminosity of the source must be in the range $10^{41} \leqslant L_p \leqslant 10^{43}$ erg s^{-1} if CR acceleration occurs inside the shell. In this case, CR in the shell of SN1987A can be deduced from the presence of gamma radiation with $E_\gamma \approx 70$–1,000 MeV (balloon observations) and $E_\gamma \geqslant 1$ TeV (ground-based Cerenkov observations). A CR luminosity $L_p \approx 10^{39}$ erg s^{-1} is needed for the latter case and $L_p \approx 10^{41}$ erg s^{-1} for the former. The source can remain active in gamma radiation until $t_a \approx 1$–3 yr, but a shorter interval is possible if $\tau_{acc} < t_a$ or mixing is small ($\eta \ll 1$). For $E_\gamma \geqslant 1$ TeV, increased mixing results in larger t_a. All this makes reasonable the idea of transferring Cerenkov telescopes from the Northern Hemisphere to the Southern. Gamma transparency of the shell sets in at a time t_γ, given by equation (13), for $E_\gamma \geqslant$ 70–1,000 MeV and equation (18) for $E_\gamma \geqslant 1$ TeV. The aforementioned observations may therefore be crucial for deciding whether or not acceleration inside supernova shells is the main mechanism of CR acceleration in the Galaxy.

Other high energy radiations discussed here demand higher, or much higher, L_p to be detectable.

If CR are accelerated at the outer shock front of supernovae, the consequent gamma radiation is expected to be too weak for detection by existing instruments.

We thank Dr V. S. Ptuskin for helpful discussions and Dr R. Staubert for informing us of and discussing the results of the balloon experiment.

Received 20 August; accepted 11 September 1987.

1. Berezinsky, V. S. & Prilutsky, O. F. *Astr. Astrophys.* **66**, 325-334 (1978).
2. Berezinsky, V. S., Ginzburg, V. L. & Prilutsky, O. F. *Sov. Astr. Lett.* **10**, 38-42, (1984).
3. Ginzburg, V. L. (ed.) *Astrophysics of Cosmic Rays* (Nauka, Moscow, 1984) (in Russian).
4. Schwartz, S. J. & Skilling. J. P. in *Proc. 17th int. Cosmic Ray Conf.* Vol. 2, 36-41 (1977).
5. Berezinsky, V. S. & Ginzburg, V. L. *Sov. Astr. Lett.* (in the press).
6. Ostriker, J. P. *Nature* **327**, 287-288 (1987).
7. Cavallo, G. & Pacini, F. *Astr. Astrophys.* **88**, 367-379 (1980).
8. Silberberg, R. & Shapiro, M. M. in *Proc. 17th int. Cosmic Ray Conf.* Vol. 2, 26-32 (1977).
9. Berezinsky, V.S., Castagnoli, C. & Galeotti, P. *Nuovo Cim.* **8C**, 185-211 (1985).
10. Gaisser, T.K., Harding, A. & Stanev, T. *Nature* **329**, 314-316 (1987).
11. Berezinsky, V.S. & Prilutsky, O. F. *Sov. Astr. Lett.* **3**, 267-271 (1977).
12. Ginzburg, V. L. & Ptuskin, V. S. *Sov. Astr. Lett.* **7**, 325-330 (1981).
13. Lagage, P. O. & Cesarsky, C. S. *Astr. Astrophys.* **147**, 127-143 (1985).
14. Stephens, S. A. in *Proc. 19th int. Cosmic Ray Conf.* Vol. 1, 201-205 (1985).
15. Grassberg, E. K., Imshennik, V. S., Nadyozhin, D. K. & Utrobin, V. P. (in preparation).
16. Litvinova. I. Yu & Nadyozhin, D. K. *Sov. Astr. Lett.* **11**, 351-358 (1985).
17. Dermer, C. D. *Astr. Astrophys.* **157**, 223-232 (1986)
18. Berezinsky, V. S. & Prilutsky, O. F. *Sov. Astr. Lett.* **3**, 153-157 (1977).
19. Gaisser, T. K. & Stanev, T. *Phys. Rev. Lett.* **58**, 1695 (1987).
20. Ciampa, D. *et al. Proc. astr. Soc. Austr.* **6**, 336-339 (1986).
21. Protheroe, R. J. & Clay, R. W. *Nature* **315**, 205-207 (1985).

Self-gravitating accretion disks in active galactic nuclei

Isaac Shlosman & Mitchell C. Begelman

Joint Institute for Laboratory Astrophysics,
University of Colorado and National Bureau of Standards, Boulder, Colorado 80309-0440, USA

Accretion onto supermassive black holes (SBHs) is widely believed to be responsible for the phenomena of active galactic nuclei (AGN). While there has been considerable theoretical work on the inner regions of such accretion flows, little is known about the outer boundaries of AGN. Except for the broad and narrow emission lines, the observational constraints are weak, and evidence linking the emission-line regions to the accretion flow is ambiguous at best. It is not even known whether the accretion flow is fuelled by mass loss from a dense cluster of stars surrounding the SBH, or enters the nucleus from the galactic interstellar medium (ISM). Here we adopt the latter view, assuming that most of the incoming fuel forms a thin accretion disk at distances of ~ 10–10^3 pc from the SBH. Such a disk must be vertically self-gravitating. We analyse some of its thermal and dynamical properties, and conclude that its energetics is likely to be dominated by backscattered AGN radiation. We also discuss the conditions under which Jeans fragmentation of such a disk into weakly interacting cloudlets can be avoided, and give a necessary condition for disk fragmentation to occur.

It has been pointed out[1-6,19] that a standard accretion disk carrying a mass flux adequate for an AGN and extrapolated to distances $\gtrsim 1$ pc from an SBH tends to become self-gravitating in the vertical direction. This leads to Jeans instability[7,8] when the mean density inside the disk exceeds the mean density associated with the external gravitational field by a small factor (2.9 for a homogeneous and isothermal keplerian disk), that is, for

$$n \gtrsim n_{\rm crit} \approx 2\times 10^7 \left(\frac{v_{\phi 2}}{r_{\rm pc}}\right)^2 \quad {\rm cm}^{-3} \qquad (1)$$

where $v_{\phi 2} \equiv v_\phi/100\ {\rm km\ s^{-1}}$ is the local keplerian speed and $r_{\rm pc}$ is the radius in pc. For an α-disk model (viscous stress $\sim \alpha p$, where p is the internal pressure), the Jeans instability sets in when the accretion rate exceeds a threshold:

$$\dot{M}_{\rm crit} \approx 5\times 10^{-4}\ \alpha T_2^{3/2}\ M_\odot \quad {\rm yr}^{-1} \qquad (2)$$

where $T = 100 T_2$ K is the equatorial temperature of the disk. The vertical component of the external gravitational field is neglected in equation (2) and further on, compared to the disk vertical gravity, which introduces an error $\sim 0(1)$. Note that $\dot{M}_{\rm crit}$ is independent of both r and the rotation curve of the disk. $\dot{M}_{\rm crit}$ is too small to supply the fuel requirements of the more luminous AGN, unless the temperature can be maintained at $\sim 10^4$ K (and $\alpha \sim 1$) or is $\alpha > 1$. While the temperature criterion is easily met (with $\alpha \lesssim 1$ at sufficiently small r, simple heating and cooling arguments suggest it is violated at $r \gtrsim 1$ pc. This is particularly true if we assume that the incoming gas has a normal interstellar dust-to-gas ratio, as the interior of the disk will then be shielded from scattered AGN hard-continuum radiation, and will probably also be rich in molecular hydrogen, an efficient coolant at low values of T_2. Except for non-local 'viscosities' (for example, external magnetic[9] or gravitational[10] torques), one can also make the case that α is unlikely to be >1 in a non-self-gravitating disk.

We are therefore forced to consider the properties of an accretion disk that is strongly self-gravitating in the vertical direction. For an α-disk the accretion rate is given by

$$\dot{M} \approx 5\times 10^{-4}\ \alpha T_2^{3/2} (n/n_{\rm crit})^{1/2}\ M_\odot \quad {\rm yr}^{-1} \qquad (3)$$

As such disks have a tendency to fragment into rings or clumps through the Jeans instability, it is important to question whether a quasi-continuous gaseous disk is a reasonable approximation here. If the gas has a relatively rigid equation of state, the instability will saturate before the clumps have contracted (due to self-gravity) by a large factor. In this case, the clumps will continue to interact strongly, leading to the continued transfer of angular momentum and the persistence of highly turbulent disk accretion. One intriguing consequence of this outcome is the possibility that the gravitational instability itself may be responsible for the viscosity[1,11]. If the clumps are able to cool on a sufficiently short timescale, however, then we would expect them to collapse to small cross-sections and to cease interacting hydrodynamically with their neighbours[12]. In this case, the gaseous disk would become what is effectively a disk of point masses, and accretion would be affected.

We can turn the above argument into a crude 'sufficient' condition for a disk *not* to fragment (or, equivalently, a necessary condition for a disk to fragment), as follows. In an accretion disk in thermal equilibrium with no external energy sources, the cooling timescale must equal the viscous dissipation timescale, $t_{\rm cool} \sim t_{\rm diss}$. In a non-self-gravitating disk $t_{\rm diss}$ is roughly a factor α^{-1} longer than the orbital time, $t_{\rm orb}$, and both timescales are much shorter than the viscous inflow time, $t_{\rm in}$. Now in a self-gravitating disk, the growth time of a Jeans instability of wavelength λ (in the plane of the disk) is roughly $t_{\rm J}(\lambda) \sim t_{\rm orb}(n_{\rm crit}/n)^{1/2}(\lambda/h)^{1/2}$, where h is the disk thickness and $h \lesssim \lambda \lesssim (n/n_{\rm crit})h$. This timescale also represents a compressional heating timescale, where the energy source is the self-gravity of the collapsing fragments. But such a disk must still satisfy $t_{\rm cool} \sim t_{\rm diss} \sim \alpha^{-1}\ t_{\rm orb}$. A sufficient condition for compressional heating to stop the collapse of fragments of wavelength λ is therefore $t_{\rm cool} > t_{\rm J}(\lambda)$, or $\alpha < (n/n_{\rm crit})^{1/2}(h/\lambda)^{1/2}$. This suggests that it is plausible for accretion to be maintained with $\alpha \gg 1$ in a highly self-gravitating disk, as proposed by Lin and Pringle[11]. But in the Lin and Pringle picture the largest unstable modes dominate the transport of angular momentum. According to our criterion, these modes may be susceptible to fragmentation if $\alpha \gtrsim 1$. If the angular momentum is transported by the shortest unstable modes, then α may be as large as $(n/n_{\rm crit})^{1/2}$ without the danger of fragmentation, although we note that this is still a factor $(n/n_{\rm crit})^{1/2}$ smaller than the value proposed by Lin and Pringle.

As we demonstrate below, it is likely that the thermal balance inside a self-gravitating AGN disk is determined not by internal viscous dissipation, but by external energy sources. In this case $t_{\rm cool} < t_{\rm diss}$, and the non-fragmentation condition on α (for angular momentum transport by the shortest modes) is more stringent, that is,

$$\alpha < \left(\frac{n}{n_{\rm crit}}\right)^{1/2} \left(\frac{t_{\rm cool}}{t_{\rm diss}}\right) \qquad (4)$$

The temperature is important in determining the mass flux that can be carried by the disk, and the relevant heating and cooling rates have a further bearing on whether the flow will resemble an accretion disk or an ensemble of weakly interacting clumps. The viscous heating rate per unit volume is

$$n^2\ \Gamma_{\rm visc} \sim 2\times 10^{-18}\ \alpha \left(\frac{n}{n_{\rm crit}}\right) \times T_2 \left(\frac{v_{\phi 2}}{r_{\rm pc}}\right)^3 \ {\rm erg\ cm^{-3}\ s^{-1}} \qquad (5)$$

A second energy source which is bound to contribute is the radiation from the AGN, backscattered onto the disk from high latitudes. Although the backscattered fraction and the degree of spectral reprocessing are highly uncertain, it is likely that the disk will be irradiated by as much as a few per cent of the AGN flux in the optical–X-ray band. The column density in the disk is $N \sim 7\times 10^{23}\ (n/n_{\rm crit})^{1/2}\ T_2^{1/2}\ (v_{\phi 2}/r_{\rm pc})\ {\rm cm}^{-2}$, which implies that all of the incident optical and ultraviolet flux will be absorbed by dust in the surface layers, out to ~ 1 kpc. For an incoming flux of $F_{\rm back} \sim 8\times 10^5\ \eta_{-2} L_{46} r_{\rm pc}^{-2}$ erg cm^{-2} s^{-1} (where L_{46} is the

AGN luminosity in units of 10^{46} erg s^{-1} and $10^{-2}\,\eta_{-2}$ is the backscattering efficiency), grains in the surface of the disk would be heated to a temperature $T_{g,\,surf} \sim 350\ (\eta_{-2}\,L_{46})^{1/4}\,r_{pc}^{-1/2}$ K. Grains with other compositions will behave in a qualitatively similar fashion. Grains in the interior of the disk will be heated by the infrared radiation which penetrates the disk, and in the absence of additional heating and cooling due to interaction with the gas, will equilibrate at some lower temperature. Hard X rays penetrating the disk will cause some ionization and direct heating of the gas, at a rate comparable to Compton heating. Adopting a backscattered fraction $\eta_X \sim 10^{-3}$ for the hard X rays and an inverse Compton temperature (for the radiation that penetrates the disk) of $10^7\,T_{IC7}$ K, we obtain a volume X-ray heating rate of

$$n^2\Gamma_X \sim 10^{-14}\left(\frac{\eta_X}{0.001}\right)L_{46}\left(\frac{n}{n_{crit}}\right) \times v_{\phi_2}^{-2}\left(\frac{v_{\phi 2}}{r_{pc}}\right)^4 T_{IC7}\quad \text{erg cm}^{-3}\,\text{s}^{-1} \qquad (6)$$

The associated primary ionization fraction is $\lesssim 10^{-4}$, which has little effect on cooling rates[14]. Note that equation (6) represents the X-ray heating rate only if the Compton y-parameter through the disk is small at the X-ray energies of interest ($y = \max(\tau_e, \tau_e^2)[(y_K T)/(m_e c^2)]$, where τ_e is the electron scattering optical depth).

Using the heating rates estimated above, we performed crude calculations of the thermal properties of a vertically self-gravitating disk, under the assumption that it carries a given mass flux $\dot{m}\,M_\odot$ yr^{-1} according to equation (3). The principal coolants at 10^2 K $\lesssim T \lesssim 10^4$ K are optically thin H_2 (ref. 15), the fine structure lines of such species as O I and C II (or C I)[14] (which may be collisionally thermalized and/or optically thick), and dust[16]. The cooling efficiency of CO and other molecular species[17] will be reduced in the high-column densities encountered in the disk. The importance of CO will rise for $\dot{m} \ll 1$ and $T_2 < 1$. We assumed that the AGN luminosity satisfies $L_{46} \sim \dot{m}/3.5$, corresponding to a mass-to-energy conversion efficiency of 0.05, took $\eta_{-2} = \eta_X/10^{-3} = T_{IC7} = 1$ and used $v_{\phi 2} = 2.5$ in several places; the models are insensitive to these assumptions. We computed heating and cooling curves for $10^{-2} \lesssim \dot{m}/\alpha \lesssim 10$ and 100 K $< T < 10^4$ K, and for the radial variable $(r_{pc}/v_{\phi 2})$ in the range 0.1–10^3. In several places we were forced to choose an independent value for α, and we chose $\alpha = 1$. Choosing a larger value of α at fixed $\dot{m}/\alpha$ increases the importance of X-ray heating and dust heating and cooling relative to viscous heating and atomic and molecular cooling. To determine the relative importance of different heating and cooling agents, we separately balanced H_2 + atomic cooling versus viscous heating; H_2 + atomic cooling versus X-ray heating; and H_2 + atomic + dust cooling versus viscous + X-ray + dust heating.

In every case we found that the temperature in the disk decreases with decreasing $\dot{m}/\alpha$ and increasing $r_{pc}/v_{\phi 2}$, and lies at $T \lesssim 1{,}500$ K for all values of the parameters. For the cases where we neglect dust, the steepness and dominance of the H_2 cooling curve at temperatures greater than a few hundred K tend to make the equilibrium temperature relatively insensitive to the heating rate in this temperature range. Thus, although the X-ray heating rate can be 2–3 orders of magnitude larger than the viscous heating rate for sufficiently large $\dot{m}/\alpha$ and small $r_{pc}/v_{\phi 2}$, the equilibrium temperature does not differ by more than a factor 2–3. At lower temperature, atomic cooling gives a relatively flat curve, and the equilibrium temperature is more sensitive to parameters. At large $r_{pc}/v_{\phi 2}$ (in the range 10–300, depending on $\dot{m}/\alpha$), the equilibrium temperature drops to <100 K.

The situation is qualitatively different when we include dust. At temperatures where thermal equilibrium is reached, dust heating and cooling are vastly more important than any other heating and cooling mechanisms. Thus, the gas must come into thermal equilibrium with the dust temperature, which in turn is set by radiative equilibrium with the ambient infrared radiation inside the disk. Radiative transfer into the disk occurs in three regimes, depending on the optical depth of the grains to the ambient radiation. Because the radiation enters symmetrically from both sides of the disk, the energy density is uniform and is given by $U_{rad} \sim F_{back}/c$. The optical depth to radiation, τ_{gr}, at temperature T_{rad} is $\tau_g(T_{rad}) \sim 10^3 \times (\dot{m}/\alpha)\,T_{rad,2}^2\,T_{g,2}^{-1}(v_{\phi 2}/r_{pc})$, where we have assumed graphite grains with a Planck-mean radiative efficiency $Q_P(T) = 6.3 \times 10^{-8}\,T^2$ (refs 13, 20). In the first regime, the disk is optically thin to the radiation emitted by the surface layer, and the grain temperature is given by the solution of the equation $\sigma T_g^4 Q_p(T_g) \sim c U_{rad} Q_p(T_{g,surf})$, where σ is the Stefan–Boltzmann constant. In the second regime, the disk is optically thick to the radiation from the surface, but thin to the radiation emitted by the grains. The radiation temperature in the interior is then determined by the condition $\tau_g(T_{rad}) \sim 1$, while T_g is determined by the equation given above, with T_{rad} replacing $T_{g,surf}$. In the third regime the grains are optically thick to their own radiation, and the interior of the disk is in local thermal equilibrium with a black-body temperature

$$T_{BB} \sim 250\ \eta_{-2}^{1/4}\,\dot{m}^{1/4}\,r_{pc}^{-1/2}\ \text{K} \qquad (7)$$

The latter regime applies for $r < 200\ \alpha^{-2/3}\,\eta_{-2}^{1/6}\,\dot{m}^{5/6}\,v_{\phi 2}^{2/3}$ pc.

Because each heating and cooling curve was computed under the assumption of constant $\dot{m}/\alpha$ given by equation (3), the density implicitly varies $\propto T^{-3}$ in each curve. As the H_2 + atomic cooling functions increase more steeply with density than either the viscous or X-ray heating curves, we expect that these thermal equilibria are stable in the temperature range under study. The case in which the gas temperature follows closely the dust temperature is also stable.

Thus, a self-gravitating accretion disk with a normal interstellar composition should remain rather cool, dusty and molecular to within 1 pc of the nucleus, even if the intrinsic viscous heating is overwhelmed by heating due to backscattered AGN radiation. This justifies the claim that an accretion disk with $\alpha \lesssim 1$ must become self-gravitating if it is to carry the mass flux required to power a luminous AGN. However, there are constraints on how large a mass flux a vertically self-gravitating disk can carry without becoming radially self-gravitating. Onset of radial self-gravitation would lead to bar instabilities and the development of large-scale spiral density waves, which could transport angular momentum on a timescale as short as the orbital time. Avoiding radial self-gravitation requires $\dot{m}/\alpha \lesssim 4 \times 10^{-2}\,v_{\phi 2}^3/T_2$, that is, $\alpha \gg 1$ for $\dot{m} \sim 0(1)$. Furthermore, the inflow time is given by $t_{in} \sim 9 \times 10^8\,\alpha^{-1}\,T_2^{-1}\,v_{\phi 2}\,r_{pc}$ yr, which is too long for a quasi-steady disk to be set up at $r \gtrsim 10$ pc, unless $\alpha > 1$.

External pressure acting on the disk places further constraints on the validity of a self-gravitating model. This pressure is not likely to be $>10\%$ of the total radiation pressure associated with the AGN luminosity in the inner kpc (although it may exceed the backscattered radiation pressure if the radiation drives an X-ray heated wind through the galaxy[18]. Thus, we have $p_{ext} \lesssim 8 \times 10^{-5}\,\dot{m} r_{pc}^{-2}$ dyn cm^{-2}. The pressure inside the disk is $p_{disk} \sim (\dot{m}/\alpha)^2\,T_2^{-2}\,v_{\phi 2}^2\,r_{pc}^{-2}$ dyn cm^{-2}. The condition $p_{ext} < p_{disk}$ is easy to satisfy if $\alpha \lesssim 100$; at much larger values of α (or $\dot{m} \ll 1$), the disk may become pressure-confined, rather than self-gravitating.

Finally, we consider the likelihood that a vertically self-gravitating disk can persist without fragmenting into weakly interacting clumps. For a disk which is dominated by external heating, the sufficient condition for *non*-fragmentation (equation (4)) can be converted to a necessary condition for fragmentation which is independent of α:

$$\frac{t_{cool}}{t_{orb}} < \left(\frac{n_{crit}}{n}\right)^{1/2} \qquad (8)$$

Particularly for disks in which the temperature is regulated by dust, condition (8) is often satisfied. Therefore, it is possible

that some of the matter in the disk will collapse and form stars. The presence of a substantial population of stars formed in the disk will undoubtedly affect both the dynamics and the energetics of the system; the nature of these effects is an important topic for further research[12].

We are grateful to Julian Krolik, Doug Lin, Dick McCray, Sterl Phinney and Mike Shull for helpful discussions, and to Doug Lin for providing us with results before publication. This work was supported in part by NASA and the NSF and grants from Rockwell International Corporation and Ball Aerospace Systems Division. M.C.B. also acknowledges support from the Alfred P. Sloan Foundation. Some of this work was done at the Institute for Theoretical Physics, University of California, Santa Barbara, which is supported in part by the NSF, supplemented by funds from NASA.

Received 16 June; accepted 12 August 1987.

1. Paczynski, B. *Acta astr.* **28**, 91-109 (1978).
2. Shields, G. & Wheeler, J. C. *Astrophys. J.* **222**, 667-674 (1978).
3. Kozlowski, M., Wiita, P. J. & Paczynski, B. *Acta astr.* **29**, 157-176 (1979).
4. Sakimoto, P. J. & Coroniti, F. V. *Astrophys. J.* **247**, 19-31 (1981).
5. Shore, S. N. & White, R. L. *Astrophys. J.* **256**, 390-396 (1982).
6. Collin-Souffrin, S. *Astr. Astrophys.* (submitted).
7. Toomre, A. *Astrophys. J.* **139**, 1217-1238 (1964).
8. Goldreich, P. & Lynden-Bell, D. *Mon. Not. R. astr. Soc.* **130**, 97-130 (1965).
9. Blandford, R. D. & Payne, D. G. *Mon. Not. R. astr. Soc.* **199**, 883-903 (1982).
10. Lin, D. N. C., Pringle, J. E. & Rees, M. J. Preprint.
11. Lin, D. N. C. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **225**, 607-613 (1987).
12. Shlosman, I. & Begelman, M. C. (in preparation).
13. Kellman, S. A. & Gaustad, J. E. *Astrophys. J.* **157**, 1465-1467 (1969).
14. Dalgarno, A. & McCray, R. A. *A. Rev. Astr. Astrophys.* **10**, 375-426 (1972).
15. Lepp, S. & Shull, J. M. *Astrophys. J.* **270**, 578-582 (1983).
16. Low, C. & Lynden-Bell, D. *Mon. Not. R. astr. Soc.* **176**, 367-390 (1976).
17. Hollenbach, D. & McKee, C. F. *Astrophys. J. Suppl.* **41**, 555-592 (1979).
18. Begelman, M. C. *Astrophys. J.* **297**, 492-506 (1985).
19. Kolykharov, P. I. & Sunyoev, R. A. *Sov. Astr. Lett.* **6**, 357-361 (1980).
20. Gilman, R. C. *Astrophys. J. Suppl.* **28**, 397-403 (1987).

Limits on quantitative information from high-resolution electron microscopy of $YBa_2Cu_3O_7$ superconductors

N. P. Huxford, D. J. Eaglesham & C. J. Humphreys

Department of Materials Science and Engineering,
University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

High-resolution transmission electron microscopy (HREM) has been used to try to characterize high-temperature superconducting ceramics on a microscopic scale. Recent studies[1,2] have concentrated on observing the ordering of oxygen vacancies in $YBa_2Cu_3O_7$, and have led to the conclusion that the presence of order, but not the degree of order, can be seen in [010] projection images under certain conditions. A related question is whether HREM images are sensitive to the oxygen concentration, that is, to the value of x in $YBa_2Cu_3O_{7-x}$. If this were the case, HREM would be a powerful way to detect local fluctuations in oxygen concentration, which may be a factor limiting the critical current. Here we present a series of image simulations in which we examine the sensitivity of HREM to both the degree of order and the oxygen stoichiometry, taking into account the uncertain nature of many of the variables in these calculations. We find a strong dependence of the simulated images on defocus, and on the Debye–Waller factor. As both of these are to some degree unknown, we conclude that although HREM can detect order, it cannot be used with confidence to quantify either the degree of ordering in the material or the oxygen stoichiometry.

HREM images depend critically not only on the structure of the material, specimen thickness, microscope parameters and image defocus, but also on the Debye-Waller factor (the extent of thermal vibration of the atoms). Previous studies[1,2,3] have not mentioned the influence of the Debye-Waller factor, nor the values used in their calculations. Here we show that images are, in fact, sensitive to the Debye-Waller factor, and also to the defocus. Simulations were carried out using a standard multislice algorithm based on the SHRLI suite[4], adapted to use a rectangular, 512×32 Fourier transform. (Parameters: voltage, 400 kV; C_s, 1 mm; Δ, 100 Å; beam divergence, 1 mrad; objective aperture, 1.2 $Å^{-1}$. The simulations are not sensitive to small changes in these parameters.)

Our calculations show that the continuous white line on the basal plane (vertical in the figures) used as the 'fingerprint' of oxygen vacancy ordering[1] is sufficiently sensitive to the defocus for quantification of the degree of ordering to be made very

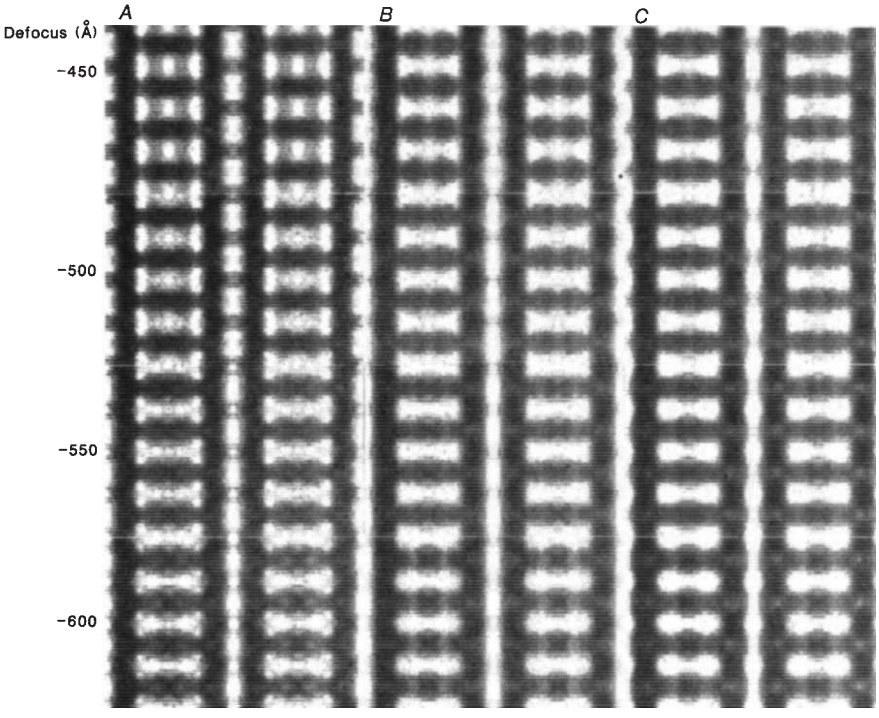


Fig. 1 Simulated [010] lattice images of $YBa_2Cu_3O_7$ for a foil thickness of 117 Å at defocus values of −450, −500, −550 and −600 Å, for *A*, complete oxygen vacancy ordering; *B*, 75% occupancy along *b*, 25% occupancy along *a*, giving a degree of ordering of 50%; *C*, complete vacancy disorder. Tetragonal unit cell parameters were used in *C*, although simulations showed the images not to be sensitive to small variations in the unit cell. The vertical white line in the centre of *A* disappears on going from *A* to *C*. The sensitivity of the details of the images to defocus illustrates the difficulty of determining the degree of ordering quantitatively.

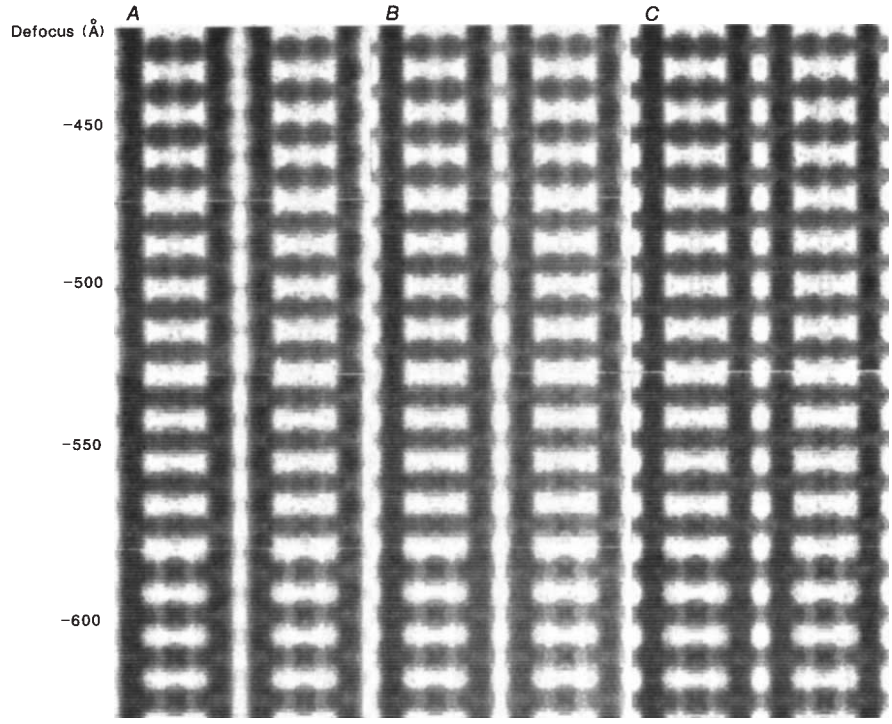


Fig. 2 Simulated [010] images of configuration *A* in Fig. 1, again at a thickness of 117 Å and for defocus values of −450, −500, −550 and −600 Å, but with different values for the Debye-Waller factors. *A*, all Debye-Waller factors set at zero; *B*, 'averaged' values of Debye-Waller factors from the literature[5-7]; *C*, the highest reported values for the Debye-Waller factors[8]. As the Debye-Waller factor is essentially an unknown in the simulations, the fact that changes in this parameter can mimic changes in oxygen ordering means that great care must be exercised in deducing the degree of ordering from HREM images.

doubtful. Figure 1 uses an 'average' value of the Debye-Waller factors found experimentally[5-7] and shows simulated images of fully ordered (Fig. 1*A*), 50% ordered (75% occupancy on the 'occupied' site and 25% occupancy on the 'vacant' site; Fig. 1*B*) and fully disordered (Fig. 1*C*) structures of $YBa_2Cu_3O_7$ at a range of defocus values (−450, −500, −550 and −600 Å), illustrating the variation in detail on the basal plane with defocus. As the precise value of the defocus is determined by fitting to image simulations, it will be difficult to quantify the degree of order with any certainty from HREM images.

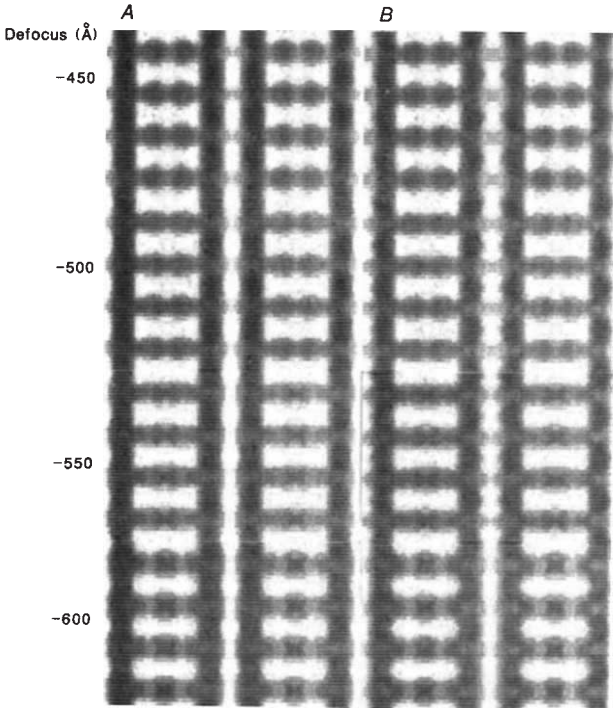


Fig. 3 Simulated [010] images for partially ordered oxygen configurations. Thickness and defocus values as in Figs 1 and 2. *A*, $YBa_2Cu_3O_{6.5}$ with the 'occupied' site half occupied and 'unoccupied' site unoccupied; *B*, $YBa_2Cu_3O_{7.5}$ with the 'occupied' site fully occupied and 'unoccupied' site half occupied. Although the images are successful in distinguishing the ordering of the oxygen vacancies, they do not appear to be sufficiently sensitive to the total oxygen content to be used in studies of stoichiometry.

Furthermore, the range of values of the Debye-Waller factor which have been deduced from neutron and X-ray diffraction studies[5-8] is wide enough both to reduce and to eliminate the 'fingerprint' for ordering if the highest values are correct. This is shown in Fig. 2, which presents simulated images of the ordered structure with Debye-Waller factors set to zero (Fig. 2*A*), set to the 'averaged' values mentioned above (Fig. 2*B*), and set to the highest values found in the literature[8] (Fig. 2*C*). Not only do the images show changes in detail, but the white line 'fingerprint' is damped out completely for the highest Debye-Waller factors. We therefore conclude that great caution is necessary in deducing the degree of oxygen ordering in $YBa_2Cu_3O_7$ unless the Debye-Waller factor can be unambiguously determined by some other technique.

Figure 3 shows further simulations, illustrating the insensitivity of images to a change of the value of x in $YBa_2Cu_3O_{7-x}$. The simulations are for $YBa_2Cu_3O_{6.5}$ with only 50% occupancy on the 'occupied' oxygen site (Fig. 3*A*) and for $YBa_2Cu_3O_{7.5}$ with 100% occupancy on the 'occupied' site and 50% occupancy on the 'vacant' site (Fig. 3*B*). Although the detailed structure of the white line on the basal plane does change, these images could not be reliably differentiated in experimental images.

We thank Drs J. C. H. Spence and A. Ourmazd for stimulating discussions on the interpretation of the image simulations.

Received 2 September; accepted 28 September 1987.

1. Ourmazd, A. & Spence, J. C. H. *Nature* **329**, 425-428 (1987).
2. Tang, D., Zhou, W. & Thomas, J. M. *Ange Wandte Chem.* (in the press).
3. Hewat, E. A., Dupuy, M., Bourret, A., Capponi, J. J. & Marezio, M. *Nature* **327**, 400-402 (1987).
4. O'Keefe, M. A., Busek, P. R. & Iijima, S. *Nature* **274**, 322-324 (1978).
5. Hewat, A. W., Capponi, J. J., Chaillout, C., Marezio, M. & Hewat, E. A. *Solid State Commun.* (in the press).
6. Capponi, J. J. *et al. Europhys. Lett.* **3**, 1301-1308 (1987).
7. Izumi, F. *et al. Jpn appl. Phys.* **26** L649-L651 (1987).
8. Okamura, F. P., Sueno, S., Nakai, I. & Ono, A. *Mater. Res. Bull.* (in the press).

Current to the ionosphere following a lightning stroke

L. C. Hale & M. E. Baginski*

Communications and Space Sciences Laboratory, Department of Electrical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Satellite[1] and rocket[2] observations of high-energy electrons precipitated from the magnetosphere associated with lightning have focused new attention on mechanisms for coupling energy to the ionosphere and magnetosphere, particularly because such events are observed in the daytime when the propagation of radio waves through the ionosphere is very difficult because of losses in the ionospheric D-region due to enhanced electrical conductivity. Electric field measurements[3,4] have indicated electrical pulses in the ionosphere following lightning strokes which are much longer in duration than predicted by electromagnetic radiation or 'relaxation time' theories. These fields can be studied theoretically using the decay of charge 'monopoles' injected into the atmospheric conductivity profile by the lightning currents. We derive here a simple analytical expression for calculating the total current waveform to the ionosphere after a lightning stroke. The validity of this expression is demonstrated by comparison with a more rigorous computer solution of Maxwell's equations. The analytic model demonstrates that the temporal variation of the current induced in the ionosphere and global circuit and the corresponding return current in the earth depends on the conductivity profile at intervening altitudes in the middle atmosphere. A conclusion is that capacitive coupling may provide tighter coupling between the lower atmosphere and the ionosphere than usually considered, in both directions, which may help to explain observations which seem to indicate that magnetospheric phenomena may in some instances trigger lightning[5,17].

The formal application of the conservative field assumption:

$$\text{curl } E = 0 \tag{1}$$

to the transient decay of fields in an ohmic medium considering only conduction and displacement currents, leads to characteristic transient solutions of the form:

$$E(\bar{x}, t) = E_0(\bar{x})e^{-t/\tau_r} \tag{2}$$

where τ_r is the 'relaxation time', at the point of observation $\bar{x}(x, y, z)$ defined as the permittivity ε divided by the conductivity σ. Probably because (2) is in fact both the general transient decay solution of Maxwell's equations in the one-dimensional case and the three-dimensional solution in special cases, it is widely believed to be generally true for cases where (1) is approximately true to a high degree of accuracy. That this is not correct can be seen by considering an initially charged capacitor C discharging into an external resistor R, where the exponential decay of the fields between the capacitor plates is determined by the time constant RC and not the dielectric relaxation time solution of (2), which would force the capacitor to discharge only into its own leakage, thus causing an error which may be many orders of magnitude. Using (1) and (2), which mathematically requires that the time-varying Maxwell current (defined as curl H) be zero everywhere, prevents the capacitive coupling of energy from low to high conductivity regions or boundaries.

C. T. R. Wilson[6] suggested that the field changes due to lightning could be accounted for by injected charges representing the uncovered charge of opposite sign to that neutralized by lightning currents. For a normal (negative) cloud-to-earth stroke or flash, a single (positive) monopole is inserted at the altitude where the charge was neutralized, the lower (negative) charge centre of the cloud (Fig. 1*a*). The recovery curve following a field change was initially thought to be due to processes inside the cloud, but Tamura[7] pointed out that the external atmospheric conductivity was also involved, using solutions of the form of (2). Subsequent workers have improved these solutions, several using the monopole decay model[8–11] to calculate transient electric field waveforms. We address here the problem of calculating the transient Maxwell current to the ionosphere after a lightning stroke, a problem which is important in the understanding of energy deposition in the atmosphere, ionosphere, and Earth.

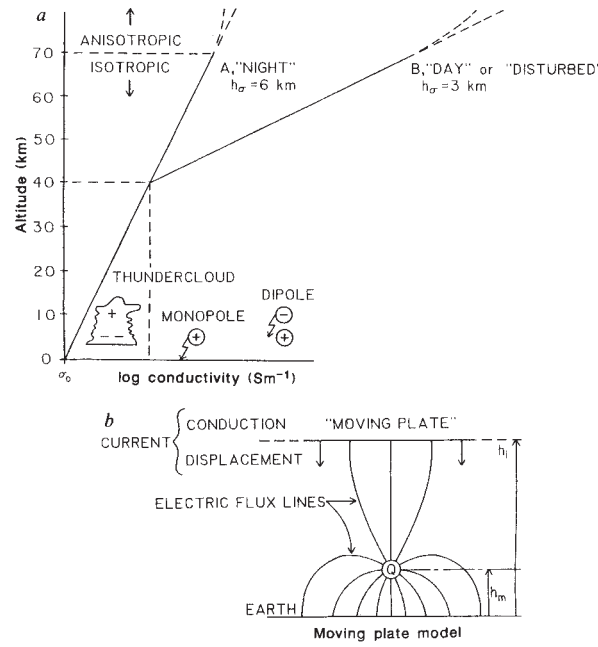


Fig. 1 *a*, Showing exponential conductivity profiles A 'night' and B 'day' or 'disturbed'. A has a single 6 km scale height (h_σ) and B a smaller (3 km) scale height above 40 km. Also indicated are the thundercloud and injected 'monopole' and 'dipole' models for calculating cloud-to-Earth and intracloud transient responses, respectively. *b*, The 'moving plate model' for calculating post-stroke transient current to ionosphere (plate) due to injected monopole Q_0, after C. and P. Greifinger[10]. The moving plate, whose altitude h_i is determined by t (after the stroke) $=\tau_r$ (relaxation time) is assumed to be the boundary between conduction and displacement current.

A paradigmatic problem is the monopole introduced into an exponential conductivity profile at $t=0$ (Fig. 1*a*, conductivity profile A). This is directly relevant to cloud-to-Earth lightning, and the superposition of two such charges can be used to treat strokes between charge centres in the atmosphere (intracloud or cloud-to-cloud). The use of a point source does not seriously impair the accuracy of the calculation at large distances. A more serious difficulty may be the fact that lightning-induced excess charges are substantial compared to the ambient ionization in even thundercloud-sized volumes of air, and hence will spread and increase the background atmospheric conductivity. Also, thundercloud conductivity may be very different from that of the surrounding air. However, these effects are not expected to be large on the time scale (up to a few seconds) considered here because most of the transient currents to the ionosphere are carried by displacement rather than conduction currents in the lower atmosphere[10]. For the purposes of solving this problem, the lower boundary provided by the Earth may be considered to be a perfect conductor.

The upper boundary is the ionosphere, which is considered

* Present address: Electrical Engineering Department, Auburn University, Auburn, Alabama 36833, USA.

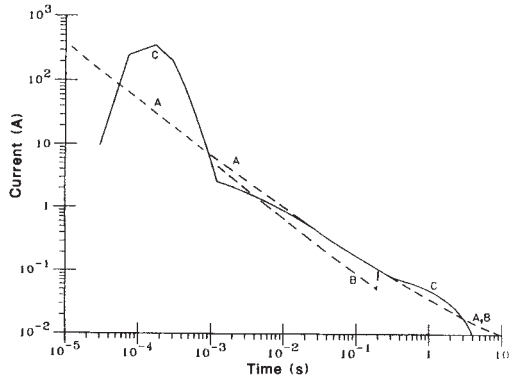


Fig. 2 Total 'Maxwell current' to ionosphere following cloud-to-Earth lightning strokes comparing analytic and computer model solutions for conditions stated. The analytic solutions, A and B, are derived from the 'moving plate' model of C. and P. Greifinger[10] following an instantaneously injected monopole Q_0. A, analytic model:

$$I_m(t) = Q_0 \left(\frac{h_m}{h_0}\right) \frac{e^{-t/\tau m}}{t\,(\ln \tau_0/t)^2}$$

B: same as A except different h_σ, σ_0, and τ_0 corresponding to different scale height above 40 km. C: Computer model, ref. 13. Conditions for models: conductivity, $\sigma(h) = \sigma_0 \varepsilon^{h/h_\sigma}$, $\sigma_0 = 5\times10^{-14}\,\mathrm{Sm}^{-1}$; $h_\sigma = 6$ km for A, C, and B (<40 km), 3 km for B (>40 km) $Q_0 = 1$C injected at $t = 0$ at $h = h_m = 6$ km. $\tau_0 = \varepsilon_0/\sigma_0 = 177$ s, $\tau_m = \varepsilon_0/\sigma_m = 65.1$ s.

Methods. The computer solution is based on the complete Maxwell's equations using an adaptation of the IMSL TWODEPEP program for solving partial differential equations to calculate the fields due to a charge injected by the more realistic Sunde double exponential model for lightning currents and hence shows source variability effects for times less than $\sim10^{-3}$ s. It uses a cylindrical volume of 80 km height and 60 km radius, with appropriate boundary conditions. The charge Q_0 is injected in finite time of order 10^{-4} s, using the Sunde[15] double exponential lightning current model into a finite (spherical gaussian) region of about 1 km radius. A and C show good agreement from milliseconds to seconds. B shows the effects of a change in conductivity scale height at 40 km, such as happens in the daytime or aurorally disturbed nighttime due to free electrons. (In reality the jump at ~0.2 s would not be expected to be so abrupt, but neither would the corresponding change in conductivity scale height at 40 km.)

here to be at the same potential as the earth. Although this is not exactly true, compared to impedances associated with thunderclouds the global circuit (a few hundred ohms in parallel with several farads) is very low. On shorter time scales (<0.1 s) the input impedance to the Earth-ionosphere 'transmission line' also defines a low-impedance external circuit.

An approximate analytic solution for the total upward current to the highly conducting ionosphere (and, because of the solenoidal nature of this current, part of the return current in the Earth) may be obtained using a viewpoint suggested by C. and P. Greifinger[10], the 'moving capacitor plate' model (Fig. 1*b*). They realized that the transient currents were mostly displacement current in the lower atmosphere and conduction current in the upper atmosphere separated by a time-dependent downward-moving boundary, which, at time t after the stroke is at an altitude h_i defined (for monotonic conductivity profiles) by:

$$t = \tau_r = \frac{\varepsilon_0}{\sigma(h_i)} \tag{4a}$$

or:

$$t = \tau_r = \frac{\varepsilon_0}{\sigma_0 e^{h_i/h_\sigma}} \tag{4b}$$

for an exponential conductivity profile, where ε_0 is free space permittivity, σ_0 is ground level atmospheric conductivity, and h_σ is the conductivity scale height. The solution for the electric field due to a point charge between two connected plates is well known, involving an infinite series of Bessel functions[12]. A much simpler property of the solution is the fraction of electric flux reaching the upper plate (see Fig. 1*b*) given by:

$$f = \frac{h_m}{h_i} \tag{5a}$$

or:

$$f = \frac{h_m}{h_\sigma \ln(\tau_0/t)} \tag{5b}$$

for an exponential conductivity profile, where h_m is the altitude of the charge Q.

Instead of the Greifinger[10] assumption of a vacuum region between the plates, we assume a small uniform conductivity σ_m corresponding to that at h_m, where a charge Q_0 is inserted at $t = 0$. This assumption has little effect for times much less than τ_m, does not change the results of (5), simplifies expression (7b) due to cancelling small terms, and accounts for the decay of the charge according to:

$$Q = Q_0\, e^{-t/\tau_m} \tag{6}$$

where $\tau_m = \varepsilon_0/\sigma_m$ is the relaxation time at the height of the charge Q. The current to the ionosphere and global circuit is given by:

$$I_m = \frac{fQ}{\tau_m} + \frac{d(fQ)}{dt} \tag{7a}$$

and, using (5b) and (6) this becomes

$$I_m = (Q_0\, e^{-t/\tau_m}) \frac{h_m}{h_\sigma} \frac{1}{t[\ln(\tau_0/t)]^2} \tag{7b}$$

for an exponential conductivity profile. The validity of these approximate expressions and of the moving plate model, which is based on replacing the diffuse conduction/displacement current boundary by a sharp moving boundary, rests on a comparison with a computer solution[13] of Maxwell's equations. This comparison is shown in Fig. 2 for a 1 C monopole injected into an exponential conductivity gradient at 6 km. The solutions appear to agree well from milliseconds to seconds. The computer solution also produced transient electric field waveforms between 20 and 50 km that were in plausible agreement with observations[3], using simultaneously measured conductivity profiles instead of the exponential profile, but insufficient diagnostics were available about magnitude and location of the lightning to allow precise comparisons.

Speed-of-light effects restrict the validity of the analytic expression, which relies on a sequence of quasistatic solutions, to times greater than a few milliseconds. If 'retarded' time at the altitude of observation is used, the computer solution should be valid to times somewhat less than a millisecond, breaking down at shorter times due to the nonspherical boundaries of the model.

Because of the long duration of these transient currents, it appears that they could be responsible for more lightning-deposited energy in the upper atmosphere than that due to electromagnetic radiation from the large lightning currents which last, usually, only for at most a few milliseconds. They may explain the observations of electric fields which persist for times much longer than the local relaxation time[3,4]. These transient waveforms have frequently been noted to be non-exponential in nature, with 'spikes' and 'long tails'. It is noted that this will describe waveforms which are relatively smooth on log-log plots, such as Fig. 2, if they are plotted on linear time and amplitude scales. However, the computer solution shows that localized $E(t)$ waveforms are quite variable in shape from point to point in the field. Because of the solenoidal nature of the Maxwell currents in the ionosphere and global circuit[14],

a current similar to that of Fig. 2 will return in the Earth in the vicinity of the storm, concentrating in high conductivity features such as rails or pipes. In the atmosphere the current would tend to focus in high conductivity structures such as rockets or conducting tethers, possibly causing corona discharge, or initiating lightning.

Equations (5a) and (7a) can be applied in their region of validity to calculate Maxwell currents for more realistic conducivity profiles. Alternatively, (7b) can be used with piecewise exponential representations of the conductivity profiles. This shows that jumps or bumps in the temporal current waveforms occur at times corresponding to the relaxation times at the altitudes of features such as 'ledges' in the conductivity profiles. The response does not depend on conductivities in the anisotropic region of the ionosphere above about 70 km for times greater than about 10^{-3} s at night and much smaller times under daytime or disturbed conditions. The most common feature is a more rapid increase in conductivity above about 40 km due to free electrons under day (or aurorally disturbed nighttime) conditions, which, using the analytic model, causes the jump at about 0.2 s shown by curve B in Fig. 2.

Intracloud strokes have been observed[3] to produce waveforms similar to those due to cloud-to-Earth strokes but of opposite polarity. Analytically, these waveforms can be studied using the decay of an injected dipole (Fig. 1*a*), employing the superposition of two monopole solutions. A more complete description of the fields, including 'steady state' currents, also requires the use of a 'meteorological generator' representing the charge separation processes in a thundercloud. This slowly varying current will not become totally dominant until at least several seconds after the lightning strokes comprising a flash, however.

The Maxwell current will be almost entirely conduction current in the high conductivity ionosphere, or at any altitude for times much greater than the local relaxation time. For the purpose of calculating more accurately the current to the ionosphere at short times, realistic source characteristics must be considered, as has been done in the computer-generated results of Fig. 2. It has been pointed out[4,10] that in the ionosphere some of the energy coupled to the ionosphere may be non-ohmic, exciting plasma waves or instabilities, but this should have little effect on the total current pulse and incident fields at the base of the ionosphere, which can then be used as the starting point for calculating more complex effects. (The computer solution indicates peak current densities of the order of 10^{-7} Am^{-2} C^{-1} incident on the anisotropic region of the ionosphere.)

The middle-atmosphere data[3] show that similar pulses of reversed polarity are generated by intracloud strokes. Since these occur globally at a much higher rate than cloud-to-Earth strokes, it may be expected that they are responsible for most of the energy coupled to the ionosphere. Cloud-to-Earth lightning, however, contributes much more directly to the (direct) current to the ionosphere, and it has been argued[3,16] that the direct current from a thundercloud to the 'global circuit' is critically dependent on the ratio of cloud-to-Earth to total lightning.

A conclusion of this paper is that simple 'capacitive' coupling of thunderstorm sources to the ionosphere, although generally neglected, is quite possible. Conversely, such capacitive coupling, via displacement current, can couple transient ionospheric and magnetospheric currents to lower altitudes. It is conceivable that these currents could trigger lightning[17], although probably only at times when breakdown was imminent, perhaps helping to explain recent observations by Armstrong[5] of lightning apparently occurring in synchronism with multiple-hop magnetospheric 'whistlers', presumably induced by lightning. The possibilities suggested by strong coupling, in both directions, between lightning and the magnetosphere open up many scenarios for further investigation.

We thank D. M. Grimes, R. H. Holzworth, and J. J. Olivero for useful discussions. This research was supported by NASA.

Received 31 December 1986; accepted 14 September 1987.

1. Voss, H. D. *et al. Nature* **312**, 740–745 (1984).
2. Goldberg, R. A., Barcus, J. R., Hale, L. C. & Curtis, S. A. *J. atmos. terr. Phys.* **48**, 293–300 (1986).
3. Hale, L. C. in *Weather and Climate Responses to Solar Variations* (ed. McCormac, B. M.) 311–321 (Colorado Associated University Press, 1983).
4. Kelley, M. C. *et al. J. geophys. Res.* **90**, 9815–9823 (1985).
5. Armstrong, W. C. *Nature* **327**, 405–407 (1987).
6. Wilson, C. T. R. *Proc. R. Soc.* A **92**, 555–574 (1916).
7. Tamura, Y., *J. Geomagn. Geoelect., Kyoto* **6**, 34–46 (1954).
8. Illingworth, A. J. *J. R. meteorol. Soc.* **98**, 604–615 (1972).
9. Dejnakarintra, M. & Park, C. G. *J. geophys. Res.* **79**, 1903–1914 (1974).
10. Greifinger, C. & Greifinger, P. *J. geophys. Res.* **81**, 2237–2247 (1976).
11. Holtzworth, R. H. & Chiu, Y. T. in *Handbook of Atmospherics*, Vol. 2 (ed. Volland, H.) 1–20 (CRC, Baton Rouge, 1982).
12. Fong, C. Y. & Kittel, C. *Am. J. Phys.* **35**, 1091–1093 (1967).
13. Baginski, M. E. thesis, Penn. State Univ. (1987).
14. Krider, E. P. & Musser, J. A. *J. geophys. Res.* **87**, 171–173 (1982).
15. Sunde, E. D. in *Earth Conduction Effects in Transmission Systems* (Dover, New York, 1968).
16. Nisbet, J. S. *J. geophys. Res.*, **90**, 9831–9844 (1986).
17. Hale, L. C. *Nature.* **329**, 769 (1987).

Phase difference between sea level and magnetic reversal rate

Stuart Gaffin*

Department of Applied Science, New York University, 26–36 Stuyvesant Street, New York, New York 10003, USA

In this letter I report a negative correlation and a phase difference of ~10 Myr between recent data on long-term eustatic sea level change[1] and geomagnetic reversal rate for the past 150 Myr. The phase difference is detectable both by visual inspection and through a correlation analysis. I also show that the same phase difference (but with a positive correlation) could theoretically have existed between this eustatic sea-level data and sea-floor creation rate. In light of the ridge-volume hypothesis of long-term eustasy, the analysis suggests the view that sea-floor creation rate (and/or subduction rate) and geomagnetic reversal rate are closely associated and perhaps synchronized in their changes.

Figure 1*a* shows the geomagnetic reversal rate for the past 150 Myr, calculated as described in the legend, using the reversal timescale used in the construction of a recently published eustatic sea-level curve[1]. An abbreviation for this data will be RF(T), T being geological time in Myr before present (BP). Figure 1*b* (open-boxed line) shows the accompanying eustatic sea level curve, also for the past 150 Myr (abbreviated as SL(T)), as determined from sequence stratigraphy on widely separated continental margins[1]. The sea level data shown is that of the 'long-term curve'[1-3], averaged over 10-Myr windows with no overlap, centred on the given year, so that it may be compared to RF(T). By visual inspection, a general negative correlation between the two data sets is apparent, with times of high sea level corresponding to low reversal rate and vice versa. Also by inspection, a phase difference between the two data sets can be seen by comparing the 100 Myr BP datum points; sea level is strongly rising while reversal rate is near its minimum; that is, the Cretaceous 'quiet zone' precedes the Cretaceous 'highstand'. In addition, a local maximum in RF(T) at 70 Myr BP precedes a local minimum in SL(T) at 60 Myr BP.

The lag is revealed more quantitatively using lagged correlation coefficients: r is computed between SL(T) and RF($T+\Delta T$) where ΔT takes sequentially the lag times ... −30, −20, −10, 0, 10, 20, 30, 40 ... Myr. Unpaired end terms are dropped off. The result of comparing RF(T) and SL(T) in this manner is shown in Fig. 2*a*. The improvement in correlation offers a quantitative estimate of the phase lag, with the peak occurring at the 10 Myr lag time, that is, when RF($T+10$) data are correlated with SL(T) data. The lag is robust at the 10 Myr averaging level. For example, shifting the averaging window onto dates centred at {5, 15, ... 155 Myr} yields a similar correlation curve. The

* Present address: Department of Geology and Geophysics, Yale University, PO Box 6666, New Haven, Connecticut 06511, USA.

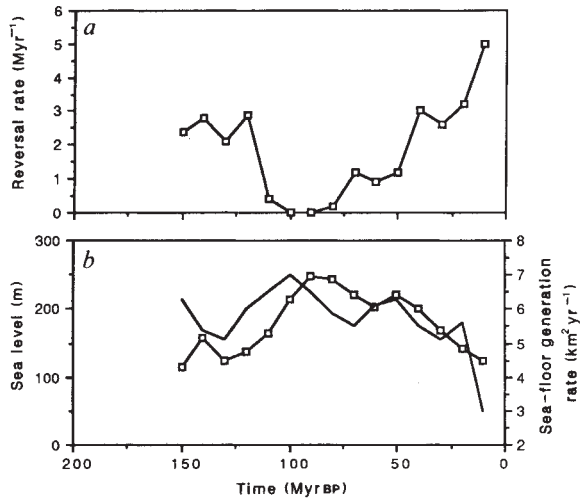


Fig. 1 *a*, Geomagnetic reversal rate post-Callovian (<155 Myr) computed in 10-Myr windows (no overlap) centred on the given year. Data were taken from ref. 1 so as to be compared to sea level changes over the same periods. *b*, Boxed line: eustatic long-term sea level changes post-Callovian (<155 Myr) relative to the present[1,2]. The data were taken from the most recent estimates (ref. 1) and placed in 10-Myr windows centred on the given year. Solid line: theoretical estimation of the corresponding sea-floor generation rates (denoted $A(T)$ in the text) that would yield, through ridge-volume dynamics, the observed sea level changes given by the boxed curve.

reversal timescale of Harland *et al.*[4], gave very similar estimates for RF(T) and the same phase lag, although any changes in reversal timescale may require changes in SL(T) that could not be made here. Previous long-term data for sea level[3], although cruder, also showed this lag when compared to the Harland *et al.*[4] reversal data. Recently the same phase lag was estimated and discussed less quantitatively by Sheridan[5] after comparing quiet zones to the sea level curve of Watts and Steckler[6]. A comparison of the two data sets using smaller averaging windows, perhaps desirable, is not as straightforward as it might seem, since the issue of the 'short-term' sea level fluctuations, their interpretation and true magnitude, then comes into play. The present study is confined to the long-term results. It is also noted that pre-Callovian reversal data (from mixed lava flows) are fundamentally less precise than the post-Callovian data (from sea-floor spreading centres) and so the analysis is restricted to the past 150 Myr.

The correlation may be essential for further insight into the mechanisms affecting either eustasy or reversals, both of which are imperfectly understood. It is reasonable to demand that the dominant mechanism controlling eustasy be plausibly linked to reversals too. A well-known mechanism, oceanic ridge volume change, certainly satisfies this demand, as this volume reflects tectonic motion and hence conditions in the mantle and perhaps in the core. Indeed, ridge-volume change is a popular explanation[7] for the broad (but not necessarily for the detailed) eustatic changes shown in Fig. 1*b*. The main pieces of quantitative evidence for this view are 'forward' ridge volume calculations[8,9] which generally show that the timescale and magnitude of the mechanism are right. However, the required seafloor creation rate estimates suffer an increasing uncertainty toward the past, with error bars large enough at 80 Myr BP that the degree of eustatic control by ridge volume remains unclear. (Similarly it is difficult to form a picture of the relationship between tectonics and reversal rate for a full 150 Myr) On the other hand, qualitative evidence from climate and geochemistry is supportive that sea level is a good proxy for global seafloor creation rate changes[10–12].

Yet it may be the phase lag which is most amenable to quantitative analysis, and it is demonstrated here that the ridge-volume mechanism offers an explanation for this lag, provided that the surficial quantity most closely associated with reversals is sea-floor creation or subduction rate (ref. 13 made an early suggestion of this association). Simply stated, ridge volume is a function of the age–depth relation for cooling oceanic lithosphere and the global rate at which new sea floor is created. If increases occur in this rate, the average age of the global ridge system decreases; in consequence, the ridge system becomes more elevated and sea level rises. However, average age varies more slowly than changes in the creation rate and a lag exists.

Quantification begins with the age–depth equation of Parsons and Sclater[14]: sea-floor depth below present sea level and above the 70-Myr sea-floor isobath is given by

$$\text{depth}(t) = 2.5 + 0.35\sqrt{t} \qquad (1)$$

where depth is in km and t is in Myr. Older sea floor also subsides (although more slowly), but its contribution will be neglected. The smallness of the resulting lag between sea level and sea-floor generation rate will justify this simplification. A key significance of equation (1) is that one need not be concerned with where past sea floor has accreted or travelled to account for it volumetrically in past ridge systems. By equation (1), the quantity $(0.35)(\sqrt{70}-\sqrt{t})$ is the difference in elevation between sea floor at age 70 Myr and seafloor at age t. An expression for ridge volume follows if, at any particular time T, this quantity is integrated over seafloor age t, weighted by the amount of sea floor of each age t:

$$\text{RV}(T) = \int_0^{70} (0.35)(\sqrt{70}-\sqrt{t})\alpha(T, t)\, dt \qquad (2)$$

where RV(T) denotes ridge volume at time T and $\alpha(T, t)$ is the differential area (per unit age) of existing sea floor of age t relative to T. It has the units of area per unit time (typically $km^2\ yr^{-1}$). Rigorously, $\alpha(T, t)$ is the amount of sea floor generated per unit time at an earlier epoch (to be identified), less the amount of sea floor lost to subduction. One estimate[15] suggests an age-dependent, subduction factor of $(1-t/180)$. This factor is near unity for small t, and does not change the phase lag estimates made below; it will therefore be ignored. In contrast, the time at which the sea floor comprising $\alpha(T, t)$ was generated is of fundamental importance. As t is measured relative to T, and T is measured relative to the present, that time must be given by the sum $(T+t)$. Letting $A(T+t)$ denote sea-floor generation rate at time $T+t$, we have

$$\alpha(T, t) \lesssim A(T+t) \qquad (3)$$

and equation (2) becomes

$$\text{RV}(T) = \int_0^{70} (0.35)(\sqrt{70}-\sqrt{t})A(T+t)\, dt \qquad (4)$$

In ref. 16 more discussion on the approximation of equation (3) can be found.

The existence of a lag is now formally apparent because ridge-volume is seen in equation (4) to depend on sea-floor generation rates at earlier times given by the argument $T+t$. To estimate its size, an inverse problem is now posed whereby changes in SL(T), as shown in Fig. 1*b*, will be used to estimate RV(T) at times in the past, and equation (4) will be inverted to estimate $A(T)$. The phase difference between $A(T)$ and SL(T) will then be studied as was done with reversal frequency. This use of sea level as a quantitative proxy for ridge volume is motivated primarily by the correlation with reversal frequency which suggests that there is a significant structure to the sea level curve worthy of investigation even with this simple model.

Details of inverting equation (4) were recently discussed[16]; only a brief description will be given here. An expansion over

time in terms of basis functions is performed on $A(T+t)$ and inserted into eq. (4) (the expansion also applies for $A(T)$). A linear system results which is solved for unknowns. For example, let

$$A(T+t)=\sum_{j=1}^{M} v_j\psi_j(T+t) \qquad (5)$$

where v_j are unknown constants and ψ_j are functions which are non-zero only over small fixed intervals of time. Inserting equation (5) into equation (4), and replacing T by a data time T_i yields:

$$\mathrm{RV}(T_i)=\sum_{j=1}^{M} v_j\int_0^{70}(0.35)(\sqrt{70}-\sqrt{t})\psi_j(T_i+t)\,\mathrm{d}t \qquad (6)$$

If N sea-level data points are used, then equation (6) represents N linear equations in M unknowns, the quantities v_j. Once these are determined, sea-floor generation rate can be computed by evaluating:

$$A(T_i)=\sum_{j=1}^{M} v_j\psi_j(T_i). \qquad (7)$$

RV(T) is related to SL(T) by a simple model of the ocean basin which assumes that other geometrical components (continental hypsometry, water volume) are fluctuating rapidly about a mean; they are taken to be fixed (of the times considered here only the late Cenozoic (0-30 Myr BP) witnessed glaciations). In this way, the total volume $TV(T)$ is defined as the integral of oceanic surface area over depth, from sea level at time T (that is, SL(T)) down to the 70-Myr sea-floor isobath (~5.4 km below present sea level, by equation (1), plus isostatic changes). By subtracting water volume (WV) from TV(T), we have RV(T) = TV(T) − WV. The ψ_j were taken as 'linear hat' functions of 20 Myr duration, with neighbours overlapping by 10 Myr; ψ_1 spanned −5 to 15 Myr BP, ψ_2 spanned 5 to 25 Myr BP, up to ψ_{25} which spanned 235 to 255 Myr BP. Sea level data were used at 10, 20, ... 240 Myr BP (the 160-240 Myr BP data are available in ref. 1 but are not shown here). Equation (6) yields a 24×25 matrix which was inverted using linear software. WV was held constant in time at 1.65×10^9 km^3, a value determined by 'shooting' the solution repetitively with trial values for WV until $A(10)$ equalled 3.0 km^2 yr^{-1}, in agreement with estimates for $A(10)$ from ref. 9. The numerical solution for $A(T)$ is shown in Fig. 1*b* (solid line), against the sea level data from which it was derived.

The existence of a lag is seen by comparing peaks and minima: a peak in $A(T)$ precedes the late Cretaceous highstand, analogous to the quiet zone, and a local minimum at $A(70)$ precedes the minimum at SL(60). Although $A(T)$ is not data, but derived from SL(T), the lag should nevertheless be analysed using the same correlation analysis (Fig. 2*b*). Similarly, $A(T)$ and RF(T) are compared in this way in Fig. 2*c*. From Fig. 2*b*, an overall 10-Myr lag emerges between SL(T) and $A(T)$. From Fig. 2*c*, it appears that a smaller phase difference exists between $A(T)$ and RF(T). The correlation peak between $A(T)$ and RF(T) is much sharper with the Harland *et al.* timescale[4], whose main difference is a longer quiet zone, but again it is unclear what modifications would be required in SL(T). A test of uniqueness for the inversion—the absence of zeros in the spectral decomposition of $(\sqrt{70}-\sqrt{t})$—shows no zeroes, and the lag is about what one expects from the half-age of the integral in equation (4), which is 17.5 Myr. The calibrating magnitude of the sea level curve is not critical as the solution will generally scale with this calibration and not affect the phase lag[16].

Correlations can at best suggest an association of $A(T)$ with RF(T) and they are used here mainly to address the question of lag time. The basic data, SL(T) and RF(T), continue to gain precision; in particular, the short-term fluctuations proposed for SL(T)[1-3] introduce an uncertainty into the ultimate shape of 10-Myr-averaged data as used here. (For that matter, though, RF(T) also shows large fluctutations on shorter timescales[17].)

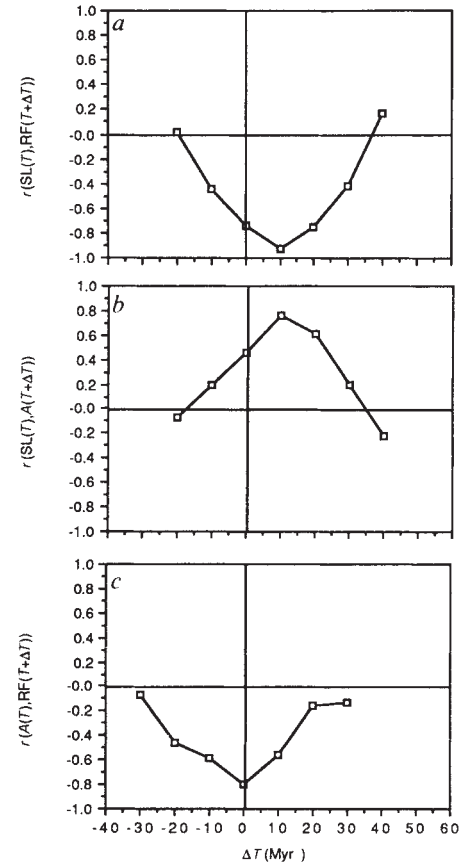


Fig. 2 Correlation coefficients between *a*, SL(T) and RF($T+\Delta T$), *b*, SL(T) and $A(T+\Delta T)$ and *c*, RF(T) and $A(T+\Delta T)$, as a function of ΔT, the lag time shown on the abcissa. T is time measured in Myr BP, increasing toward the past. Unpaired end terms in the correlations were neglected.

Similarly, the inversion is being used to address the same question of lag time, and the results for $A(T)$ should be treated as illustrative. (The inversion is generally sensitive to changes in the sea level data because it removes the smoothing of the integral in equation (4).) Nevertheless, a comparison with the integrated plate data of refs 8, 9 and 18, available for the past 80 Myr, shows some similarities: there are minima near 70 and 40 Myr BP. As well as the connection of $A(T)$ to eustasy, recent models of the carbonate-silicate cycle suggest that it may be the dominant control on long-term atmospheric carbon dioxide and climate[18]. But uncertainty in the true variation of $A(T)$ effectively reduces the model to a sensitivity study of this central quantity. If a link with reversal rate exists then it may be exploitable in different ways. An outstanding question is how $A(T)$ and reversal rate might be negatively correlated and tightly synchronized. Thermal couplings between the mantle and core, and recent views on the reversal process[19-22] can be looked at in the light of the evidence presented. A theoretical argument either supporting or denying a synchronization between the outer core and upper mantle would be valuable.

I thank B. Haq for a composite data chart and for a preprint, M. Rampino for early remarks on the negative correlation, and M. Hoffert for financial support.

Received 18 March; accepted 28 July 1987.

1. Haq, B. U., Hardenbol, J. & Vail, P. R. *Science* **235**, 1156-1166 (1987).
2. Vail, P. R., Hardenbol, J. & Todd, R. G. *Mem. Am. Ass. Petrol. Geol.* **36**, 129-144 (1985).
3. Vail, P. R., Mitchum, R. M. & Thompson, S. *Mem. Am. Ass. Petrol. Geol.* **26**, 83-97 (1977).
4. Harland, W. B. *et al. A Geologic Time Scale* (Cambridge University Press, 1982).
5. Sheridan, R. E. *Paleoceanography* **2**, 97-118 (1987).
6. Watts, A. B. & Steckler, M. S. in *Deep Sea Drilling Results in the Atlantic Ocean: Continental Margins and Paleoenvironment, Maurice Ewing Ser.* Vol. 3 (eds M. Talwani, W. W. Hay, W. B. F. Ryan) 218-234 (Washington, DC, 1979).

7. Donovan, D. T., Jones, E. J. W. *J. geol. Soc., Lond.* **136**, 187-192 (1979).
8. Pitman, W. C. *Geol. Soc. Am. Bull.* **89**, 1389-1403 (1978).
9. Kominz, M. A. *Mem. Am. Ass. Petrol. Geol.* **36**, 109-127 (1984).
10. Mackenzie, F. T., Pigott, J. D. *J. geol. Soc., Lond.* **138**, 183-196 (1981).
11. Tissot, B. *Nature* **277**, 463-465 (1979).
12. Sanders, P. *Nature* **305**, 19-22 (1983).
13. Vogt, P. R. *Earth planet. Sci. Lett.* **25**, 313-321 (1975).
14. Parsons, B., Sclater, J. G. *J. geophys. Res.* **82**, 803-827 (1977).
15. Parsons, B. *J. geophys. Res.* **87**, 289-302 (1982).
16. Gaffin, S. *Am. J. Sci.* **287**, 596-611 (1987).
17. Jacobs, J. A. *Q. Jl R. astr. Soc.* **27**, 328-350 (1986).
18. Berner, R., Lasaga, A. C. & Garrels, R. M. *Am. J. Sci.* **283**, 641-683 (1983).
19. McFadden, P. L. & Merrill, R. T. *J. geophys. Res.* **89**, 3354-3362 (1984).
20. Jones, G. M. *J. geophys. Res.* **82**, 1703-1709 (1977).
21. Gubbins, D. *Nature* **326**, 167-170 (1987).
22. Gubbins, D. & Richards, M. *Geophys. Res. Lett.* **13**, 1521-1524 (1986).

Proposed synthetic zeolite ECR-1 structure gives a new zeolite framework topology

M. E. Leonowicz* & D. E. W. Vaughan

Exxon Research and Engineering Company, Annandale, New Jersey 08801, USA

Zeolites dominate petroleum and new petrochemical processing, and any new large-pore zeolite is of major significance. Here we propose a structure for the recently synthesized zeolite phase ECR-1 (ref. 1). The proposed model is an intimate twin of the mordenite[2] and mazzite[3] structures, constructed by interposing layers of mordenite-like sheets between layers of mazzite-like cages. Although there is an ambiguity between two ways the layers are interconnected, this layer-twin model explains the results of sorption and crystallization experiments, the spectroscopic data, X-ray and electron diffraction, and high-resolution transmission electron microscopy (HRTEM) lattice imaging results. The ECR-1 channel system is one-dimensional, with 12-membered-ring openings intermediate in size between those of the parent materials. The structure solution confirms the power of HRTEM lattice imaging in helping to solve the structures of microcrystalline materials[4,5], and the intergrowth of ECR-1 and mazzite observed in the images suggests that a new subfamily of zeolite materials based on differing sequences of the two structural layers can be synthesized.

ECR-1 is synthesized with a typical Si/Al ratio of 3.6 using the bis-(2-hydroxyethyl)-dimethyl-ammonium cation in sodium silicate/sodium aluminate solutions at 150 °C. The higher Si/Al end of the crystallization field is bounded by synthetic mordenite. Lower temperature (~100 °C) syntheses give faujasite and, in the presence of trace TMA impurities, a mixture of mazzite and faujasite[6]. The ECR-1 infrared spectrum shows similarities with those of both mordenite and mazzite and has peaks at ~1,210 and 1,235 cm^{-1} in a region that has been assigned[7] to 5-rings in zeolites. The morphology of ECR-1 is lathlike, with lengths of several microns and widths of several hundred Ångstroms or less, as shown in Fig. 1.

Using lattice symmetry and approximate cell lengths from selected area electron diffraction patterns (EDP), the ECR-1 X-ray powder pattern was successfully indexed with a primitive orthorhombic unit cell $a = 7.310(4)$ Å, $b = 18.144(6)$ Å and $c = 26.31(1)$ Å, the precise values determined from least-squares refinement of 49 well-defined peak positions. While evidence for systematic absences in the X-ray pattern was inconclusive due to disparate cell lengths, some hk0 EDPs clearly indicated an *n*-glide plane perpendicular to the 26.3 Å *c* axis. ECR-1 has the diffraction symmetry P**n, for which the only centrosymmetric space group possibility is Pmmn (D_{2h}^{13}, no. 59). The EDPs and HRTEM lattice images were obtained using a Philips 420ST electron microscope operated at 120 kV with a calculated 3.0 Å

Fig. 1 Scanning electron micrograph of ECR-1 showing "wheat-sheaf" bundles of crystals <0.1 μm in diameter and 8 μm long. Scale bars, 10 μm (top) and 1 μm.

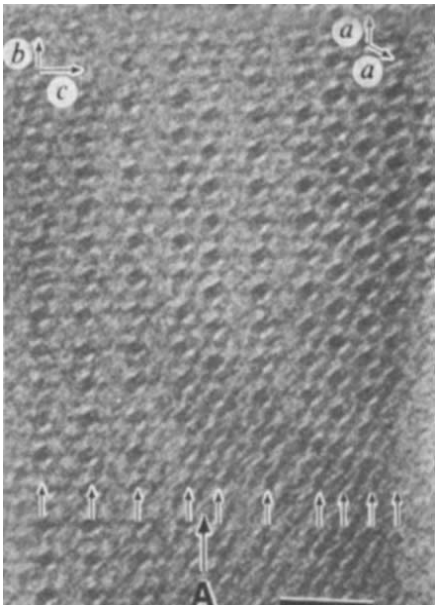


Fig. 2 High-resolution transmission electron micrograph, lattice image of ECR-1 *bc* plane showing syntactic intergrowth with a hexagonal phase to the right. Orientation of ECR-1 *b* and *c* axes and the hexagonal *a* axis is indicated. Arrows indicate layers of structure common to both materials. The feature indicated by A is a twin plane; Fig. 4 illustrates the nature of this feature in terms of the model structure.

* Present address: Mobil Research and Development Co., Paulsboro, New Jersey 08066, USA.

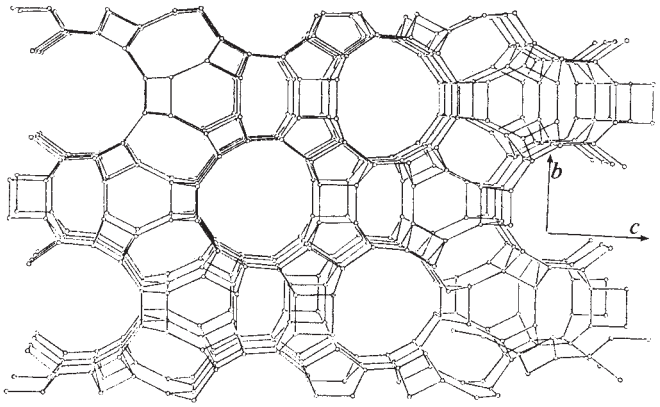


Fig. 3 Model structure proposed for the zeolite ECR-1, as viewed down the lattice *a* direction. Only tetrahedral atoms are plotted for clarity. This model shows 4,6-ring connectivity between mordenite and mazzite sheets. The alternative 5-ring connectivity model is constructed by shifting the layers by $a/2$ relative to each other. The 12-ring channels in ECR-1 shown are unevenly spaced along the 26.3-Å *c* direction.

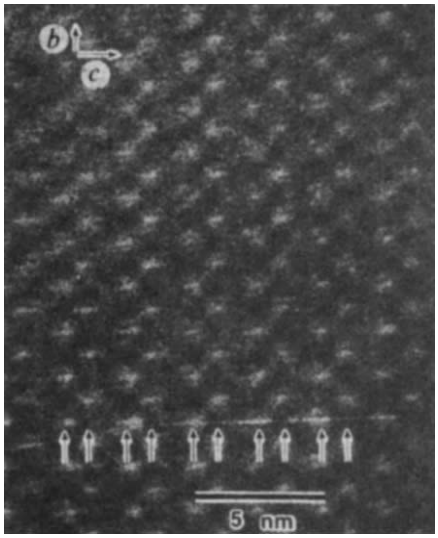


Fig. 5 High-resolution transmission electron micrograph lattice image of ECR-1 *bc* plane showing double rows of white features running in the *b* direction, which can be interpreted as the 12-ring channels in ECR-1 that are parallel to the *a* direction. The spacings between rows of channels correspond to that predicted by the model.

point-to-point resolution at optimum defocus, with lattice images limited to 2.4-Å resolution by the objective aperture.

The large unit cell ($V = 3{,}490$ Å^3) makes construction of model structures for ECR-1 difficult despite lattice constants 7.3 Å and 18.1 Å similar to those of other zeolites[8]. ECR-1 lattice images were generally uninformative until the *bc* plane (18 Å × 26 Å) was imaged in samples embedded and ultramicrotomed to reveal the plane perpendicular to the lath axis (the 7.3 Å *a* direction). An important clue to the structure is shown in Fig. 2, a *bc* plane lattice image wherein a twinning mode of the material is observed. The twin plane is the ECR-1 *ab* plane (7.3 Å × 18 Å) and the twin phase has apparent hexagonal symmetry with cell lengths $a' \simeq 18$ Å and $c' \simeq 7.3$ Å. Furthermore, careful examination reveals a close resemblance between the structural layers composing the hexagonal $a'c'$ planes and corresponding layers in the orthorhombic ECR-1 *ab* plane, suggesting that a common layer of structure exists in both materials. The orthorhombic symmetry and larger repeat distance normal to the layer would be achieved in ECR-1 by interconnecting these common layers through some other structural subunit.

There are three well-known zeolite structures with hexagonal symmetry and essentially identical lattice constants $a \simeq 18.4$ Å and $c \simeq 7.5$ Å: Linde type L[9], mazzite[3] and the hypothetical structure proposed for zeolite Omega[10]. Only mazzite, however, contains a sheet of structure compatible with an *n*-glide symmetry operator; the ~15.5-Å thick layer, consisting of gmelinite cage columns connected by 5-rings, is well described by the two-sided plane group[11] pmmn with repeat distances ~18.4 Å and ~7.5 Å within the layer. The ECR-1 structure model resulted from the realization that mordenite[3] contained a ~10.5 Å thick layer (26–15.5 Å) with the same symmetry and repeat distances and that tetrahedral network bonding demands could be fully satisfied between these two structural entities. The model structure shown in Fig. 3 can be viewed as an intimate twin of these two materials, a regular recurrent intergrowth[14] of mordenite-like and mazzite-like sheets forming a new zeolite. The presence of mazzite layers in ECR-1 explains its ability to twin syntactically with the hexagonal phase; Fig. 4 is a detailed structural illustration of the twin feature in Fig. 2.

Within the Pmmn symmetry restrictions and the constraints of three-dimensional network bonding, mazzite and mordenite layers can be interconnected in two ways: either through 5-ring chains (model A) reminiscent of the native structures, or through chains of alternating 4- and 6-rings (model B); the connectivity

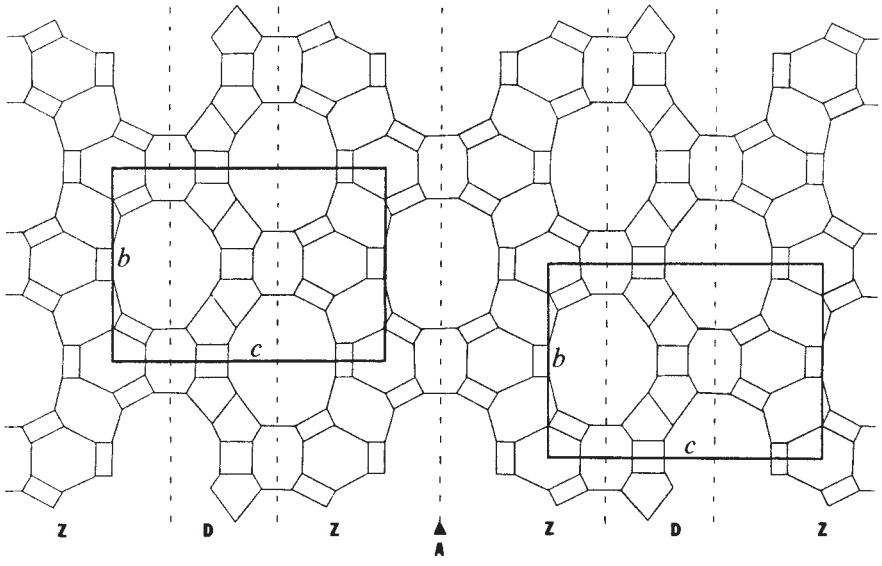


Fig. 4 Structural model for the twin feature A indicated in Fig. 2. The structure is viewed from the ECR-1 *a* direction and the *b* and *c* directions of the unit cell are indicated. Dashed lines indicate boundaries between mazzite layers (marked Z) and mordenite layers (marked D). The twin plane marked A corresponds to feature A in Fig. 2.

depends on the relative position of the layers in the 7.3 Å direction, with an $a/2$ shift between layers converting one model into the other. DLS refinement[12] of either model using ideal distances (T-0, 1.638 Å; T-T, 3.13 Å) adjusted[13] to the optimum Si/Al ratio gave essentially identical figures-of-merit[12] $R = 0.003$ and $\sigma = 0.014$; these values indicate excellent agreement with symmetry constraints and experimentally determined lattice constants. Elucidation of the precise interlayer connectivity will require full Rietveld refinement of framework atoms and channel contents but, while the calculated framework-only powder patterns for the models are very similar, examination of regions in which substantial intensity differences occur suggests that the connectivity between layers may be through alternating 4- and 6-rings, favouring model B.

Further evidence supporting the layer-twin model is shown in Fig. 5. This *bc*-plane lattice image shows double rows of white features in the 18 Å direction of the lattice, which can be interpreted as the ECR-1 12-ring channels parallel to the unit cell *a* direction. The spacing between features within the double rows is ~10 Å, while the spacing between features in adjacent double rows is ~15 Å, as measured parallel to the 26-Å repeat direction. The channels in the ECR-1 model shown in Fig. 3 are unevenly spaced, being ~10.6 Å apart across mordenite sheets and ~15.7 Å apart across mazzite sheets, as measured centre-to-centre parallel to the 26.3 Å *c* axis. The correspondence between the features observed in Fig. 5 and the predictions from the model supports the layer-twin hypothesis.

The large family of possible zeolite structures that can be constructed by interconnecting sheets of mordenite, mazzite and omega will be discussed elsewhere.

Received 17 July; accepted 1 September 1987.

1. Vaughan, D. E. W. & Strohmaier, K. G. *Proc. 7th Int. Zeolite Conf.* 207 (eds Morakami, Y., Iijima, A. & Ward, J. W.) (Kodansha/Elsevier, 1986).
2. Meier, W. M., *Z. Kristallogr.* **115**, 439 (1961).
3. Galli, E. *Crystallogr. Struct. Commun.* **3**, 339 (1974).
4. Barri, S. A. I., Smith, G. W., White, D. & Young, D. *Nature* **312**, 533–534 (1984).
5. Millward, G. R., Ramdas, S. & Thomas, J. M. *Proc. R. Soc.* **A399**, 57–71 (1985).
6. Vaughan, D. E. W. *Discussions 7th Int. Zeolite Conf.* 24 (ed. Tominaga, H.) (Japan Ass. Zeolites, 1986).
7. Flanigen, E. M. in *Zeolite Chemistry and Catalysis* ACS Monogr. 171, (ed. Rabo, J. A.) 80–117 (Am. Chem. Soc., Washington DC, 1976).
8. See, E.G., Meier, W. M. & Olson, D. H. *Atlas of Zeolite Structure Types* (Druck & Verlog, Zurich, 1978).
9. Barrer, R. M. & Villiger, H. *Z. Kristallogr.* **128**, 352 (1969).
10. Barrer, R. M. & Villiger, H. *Chem. Commun.* 659 (1969).
11. Weber, L. *Z. Kristallogr.* **70**, 309 (1929).
12. Baerlocher, Ch., Hepp, A. & Meier, W. H. *DLS-76, A Program for the Simulation of Crystal Structures by Geometric Refinement* (Inst. Crystallogr. Petrogr., Zurich 1978).
13. Meier, W. M. in *Molecular Sieves*, Adv. Chem. Ser. No. 121, (eds Meier, W. M. & Uytterhoeven, J. B.) 39–51 (Am. Chem. Soc., Washington DC, 1973).
14. Rao, C. N. R. & Thomas, J. M. *Acc. Chem. Res.*, **18**, 113–118 (1985).

Direct determination of the intramolecular O–D distance in ice Ih and Ic by neutron diffraction

M. Antonio Floriano*†, D. D. Klug*, Edward Whalley*, E. C. Svensson‡, V. F. Sears‡ & E. D. Hallman‡§

* Division of Chemistry, National Research Council of Canada, Ottawa, K1A 0R6, Canada

‡ Atomic Energy of Canada Limited, Chalk River Nuclear Laboratories, Chalk River, Ontario, K0J 1J0, Canada

The mean apparent intramolecular O–D distance in ice Ih, as determined by neutron diffraction using standard methods, is well established[1–5] as 1.01 or 1.02 ± ~0.01 Å. It is enormously greater than the O–D distance in the vapour, which is[6] 0.9732 Å, and Whalley[7] has argued that it cannot be the actual distance, and that the actual distance must be about 0.98 Å. The difference between these distances is caused by the orientational disorder of the water molecules[5]. Kuhs and Lehmann have combined accurate neutron single-crystal diffraction with thermal amplitudes obtained from vibrational spectroscopy, or by assuming that the thermal atomic displacements are purely gaussian, and have concluded that the actual O–D distance in ice Ih is 0.983 ± 0.005 and 0.979 ± 0.005 Å, respectively, from the two techniques[5]. This work is certainly a *tour de force*, but the method is not entirely direct. This letter describes a new method of determining actual interatomic distances in disordered crystals directly from diffraction measurements alone and its application to ice Ih and Ic.

The method is to convert the powder diffraction pattern of the disordered crystal, including both the Bragg peaks and the diffuse scattering, to the static structure factor, and to Fourier transform the structure factor to obtain an average neutron-weighted pair distribution function. The actual interatomic distributions averaged over all equivalent pairs of atoms are then obtained. The powder diffraction pattern of crystals has often been Fourier transformed in order to help interpret the pair distribution function of amorphous forms of the same substance (ref. 8 is one example), but the technique seems not to have been used previously to determine directly the actual interatomic distances in disordered crystals.

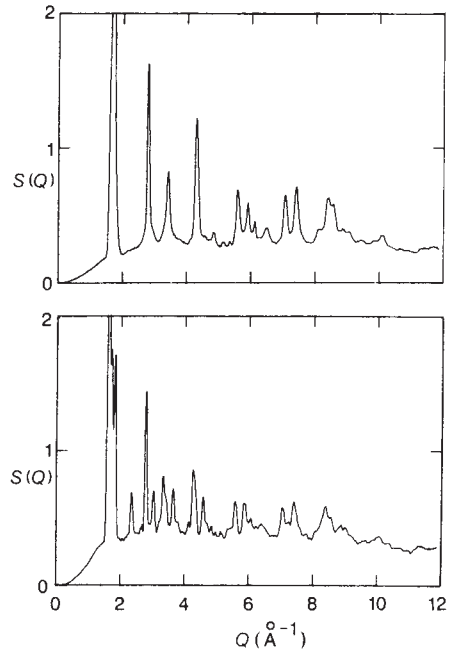


Fig. 1 Neutron static structure factor $S(Q)$, as a function of the momentum transfer Q, of the sample of polycrystalline ice Ic that had been made by heating high-density amorphous ice to 172 K, in the upper frame, and of the sample of ice Ih that was obtained by heating to 240 K the ice Ic that was made at 197 K, in the lower frame.

Polycrystalline samples of D_2O ice Ic and Ih were made as part of a neutron-diffraction study of high-density amorphous ice[9]. The high-density amorph was made as a wafer, which was ~1.5 mm thick and 69.5 mm in diameter, by sandwiching a sample of D_2O ice Ih between steel plates lined with Teflon sheets and holding it in place by cadmium and beryllium-copper

† Present address: Dept of hemistry, Purdue University, West Lafayette, Indiana, 47907, USA.
§ Permanent address, Dept of Physics & Astronomy, Laurentian University, Sudbury, Ontario P3E 2C8, Canada.

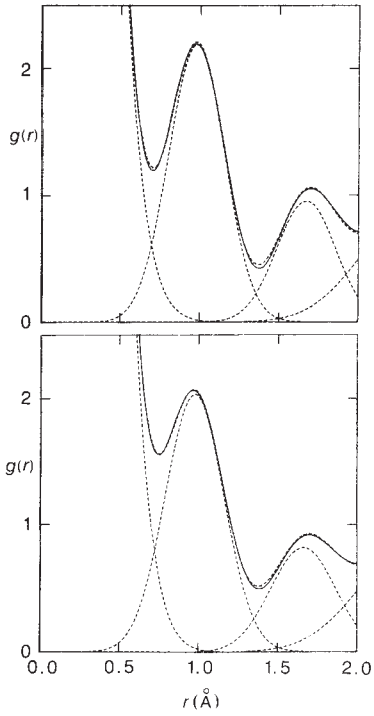


Fig. 2 Pair distribution function $g(r)$, where r is the distance from the central atom, of polycrystalline ice Ic, in the upper frame, and of ice Ih, in the lower frame, both drawn as solid lines. They were obtained from the corresponding structure factors in Fig. 1. The four gaussian components that were fitted to the distribution function as described in the text are drawn as dashed lines, and the sum of the gaussians is drawn as the dashed line that fits the distribution function closely.

rings. The sandwich was placed in a piston-cylinder apparatus, which was then cooled to 77 K by liquid nitrogen and was placed in a hydraulic press having a capacity of 1,000-tonnes-weight. The sample was squeezed to ~13 kbar[10] and the displacement of the pressurizing piston was measured to a precision of ±1 μm with an Ames dial gauge. When the transformation was complete, the press load was reduced to zero and the sample was removed from the pressure vessel under liquid nitrogen and was then sandwiched between two discs of single-crystal silicon. The sandwich was placed in an aluminium holder covered with a cadmium mask in order to restrict to a 50 mm × 50 mm square the part of the sample that scatters neutrons. The assembly was mounted at 77 K on the tail of a Displex closed-cycle refrigerator, was surrounded by an aluminium radiation shield, and then cooled to 17 K. The temperature was measured to ±1 K by two silicon diodes mounted on the top and bottom parts of the holder. The samples were converted to ice Ic (after a series of diffraction experiments that are described in ref. 9) by heating one to 172 K and one to 197 K for 5 min. The diffraction patterns of two samples of ice Ic were measured as described below, and the pattern of the sample made at 197 K was measured again after it had been converted to ice Ih by heating it to 240 K for 5 min[11].

The neutron-diffraction experiments were done on the N5 spectrometer at the NRU reactor at Chalk River, using a beam of either 0.8979-Å neutrons from beryllium (110) planes or 0.8984-Å neutrons from germanium (331) planes. During all measurements, the plane of the sample bisected the angle between the incident and scattered beams. The beam was collimated to 0.55 and 0.59° before and after the sample respectively, and the scattered intensity was measured at scattering angles of 2 to 45° at intervals of 0.2° and at scattering angles of 45 to 116° at intervals of 0.5°, the maximum momentum transfer being 11.9 $Å^{-1}$. The principal peak had a height of ~40,000 counts, and the intensity from the empty vessel was 100–500 counts except below 5°, where the number of fast neutrons rose rapidly, and in the powder peaks from the radiation shield. The scattering by the empty vessel was multiplied by a shielding factor of ~0.9, which was chosen so that when it was subtracted from the spectrum with the ice present it removed the aluminium powder peaks. All three samples gave good powder patterns.

High-density amorphous ice made at 77 K has always contained a few small unchanged crystals of ice Ih[10], and an earlier sample[10] that was kept at 175 K for 20 min transformed completely to ice Ih by growing onto these crystals. The shapes and relative heights of the powder peaks of the two samples of ice Ic, as plotted on a large graph, were the same, and the relative weights of the peaks in the sample of ice Ih agreed with a published pattern[12], both within the precision of the experiments. The structure factors are, therefore, not significantly affected by any large single crystals of ice Ih that may have been present.

The data were corrected for absorption, incoherent scattering, and multiple scattering, and for inelastic recoil scattering using the usual Placzek correction[13]. They were then multiplied by a constant to make the static structure factor $S(Q)$ at large momentum transfer Q tend to the theoretical value in the limit of infinite momentum transfer, which is $\sum_\alpha b_\alpha^2/[\sum_\alpha b_\alpha]^2$, where b_α is the coherent scattering length of nucleus α and the sum is over all nuclei in the water molecule. For D_2O water, $S(\infty) = 0.334$. The structure factor at low momentum transfer was extrapolated smoothly to the theoretical value, which is $k_B T(\partial\rho/\partial p)_T$ where k_B is the Boltzmann constant, T the temperature, ρ the density, and p the pressure, and which is about zero at $Q = 0$. The density of both forms of ice is 0.0312 molecule $Å^{-3}$. The structure factors so obtained of the sample of ice Ic that had been heated to 172 K and of the sample of Ih are plotted in Fig. 1.

The structure factor is related to the pair distribution function $g(r)$, where r is the distance from the central nucleus, by the equation

$$g(r) - 1 = \frac{1}{2\pi^2 \rho r} \int_0^\infty Q[S(Q) - S(\infty)] \sin Qr \, dQ \qquad (1)$$

where ρ is the number density of molecules. Numerical integrations were done by the method of overlapping parabolas[14] and, because the experimental range of momentum transfer is limited, Lorch's truncation method[15] was used to eliminate spurious ripples. The resulting pair distribution functions of the samples of ice Ic and Ih represented in Fig. 1 are plotted at small distances in Fig. 2. The pair distribution function is the weighted probability that, if there is a particular atom at the origin, there is another atom at distance r from the origin. The origin is placed on each atom in the sample in turn and the average taken over all atoms. The weights of the probabilities of the pairs D--D, D--O, and O--O depend on the neutron scattering lengths, and are 0.49, 0.42, and 0.09 respectively. Thus, the D--D and D--O distributions have nearly the same weight (but may have different multiplicity depending on the particular pairs) and the O--O distribution is hardly detectable. Therefore, neutron diffraction directly determines the actual intramolecular O-D distance if the diffraction pattern is analysed in this way.

The divergence of the pair distribution at low distance is spurious[13] and is caused by errors in the structure factor that are almost independent of the momentum transfer. The peak at ~0.97 Å in Fig. 2 represents the O-D distribution. Its position is slightly affected by the spurious divergence near zero radius and by the peak at ~1.7 Å, which is caused by overlapping intramolecular D-D and intermolecular hydrogen-bonded O--D distributions. To accurately determine the position of the peak near 0.97 Å, a sum of gaussians each having the form

$$g_i(r) = \frac{w_i}{\sqrt{2\pi}\,\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{r - r_{oi}}{\sigma_i}\right)^2\right] \qquad (2)$$

Table 1 Parameters of the four gaussians, as defined by equation (2) with $i = 1$–4, that were fitted to the first four features in the pair distribution of ice Ic and Ih at 17 K

	r_{oi}(Å)			w_i			σ_i/Å		
i	Ic(172 K)*	Ic(197 K)	Ih(240 K)	Ic	Ic	Ih	Ic	Ic	Ih
1	0	0	0	14.989	35.247	24.859	0.2586	0.2388	0.2604
2	0.978	0.973	0.976	0.9640	0.8959	0.9400	0.1757	0.1806	0.1845
3	1.67	1.64	1.66	0.4422	0.4652	0.4195	0.1857	0.2236	0.2041
4	2.32	2.32	2.32	0.6579	0.5744	0.6309	0.3162	0.3015	0.3227

r_{oi} is the distance of the i-th maximum, w_i its weight, and σ_i its standard deviation.
* The samples were made by heating the high-density amorph to the temperatures shown in parentheses.

where w_i and σ_i are respectively the weight and standard deviation of the pair distribution function g_i of the i-th peak and r_{oi} is the distance of the i-th maximum, was fitted to the peaks centered at zero, ~1, ~1.7, and ~2.3 Å using a nonlinear regression. The divergence at small distances is not a gaussian but approximates one very well in the tail region, and was therefore fitted to a gaussian in the range of distance where its value was less than ~6. Although the peak centred at ~1.7 Å contains both the intramolecular D–D and the hydrogen-bonded O--D correlations, fits with one or with two gaussians did not differ significantly and so only one gaussian was used in the final fit. The parameters of the fitted gaussians are listed in Table 1, and each curve and the sum of the four curves are compared with the experimental distribution in Fig. 2. The area of the O–D peak when weighted by $4\pi\rho r^2$, that is $4\pi\rho \int r^2 g(r)\,dr$, should be 0.56, and the experimental areas are 0.38, 0.35, and 0.37 for the three samples in the order they occur in Table 1. The area of the combined D–D and O--D peaks should be 0.890, and the experimental areas are 0.49, 0.52, and 0.46 for the same three samples. The experimental areas of both peaks are about 0.7 of the expected areas, probably because the structure factor is inaccurately normalized. The inaccurate normalization will not, of course, affect the peak distances.

The O–D distances in the two samples of ice Ic as determined in this way are 0.978 ± 0.002 and 0.973 ± 0.002 Å, and that of the sample of ice Ih is 0.976 ± 0.002 Å. According to Powles[16], when interference terms in the inelastic correction, which are usually neglected, are allowed for, the apparent O–D distance in liquid water increases by about 0.009 Å. If the same correction applies to the O–D distance in ice Ic and Ih, the distance in ice Ic becomes 0.987 and 0.982 Å, and in ice Ih it becomes 0.985 Å. The distances in ice Ic and Ih do not differ significantly and the average is 0.985 Å with an estimated uncertainty of 0.006 or 0.007 Å. Kuhs and Lehmann's[5] values of 0.983 ± 0.005 and 0.979 ± 0.005 Å and our value agree well. As the mean O–D distance in the vapour is 0.9732 Å (ref. 6), the O–D bond is lengthened by only 0.012 Å, with an uncertainty of 0.006 or 0.007 Å, when the vapour is condensed to ice I. An earlier suggestion[17] for the origin of the large decrease in the O–H or O–D stretching frequency is, therefore, not correct. It appears that hydrogen bonding reduces the force constant not only by stretching the O–H bond, but also by weakening it, presumably by removing electrons from it.

According to current literature[5], the actual O--O distance in ice I is also uncertain. The reason for this is that the crystallographic distance is the distance between mean positions as determined by averaging over all unit cells, and not the distance between the actual positions of the oxygen atoms. However, for perfect tetrahedral coordination, the sum of the lengths of the bonds is, by symmetry, invariant to small displacements of the atoms, and so small displacements from the ideal structure do not change the average bond length to first order. Consequently, the crystallographic bond length, which is the distance between average positions, should be the same as the mean of the actual bond lengths. The crystallographic O–O distance is[4] 2.751 Å. The actual distance, as determined by Fourier transforming the x-ray structure factor from reciprocal space to direct space[8] is 2.751 ± 0.002 Å. The mean O--O distances as determined by the two techniques are the same within the experimental precision, as they should be.

The technique described here—to Fourier transform the complete static structure factor as determined by neutron or x-ray diffraction—can be used to determine actual interatomic distances in all disordered crystals, including crystals having rapidly reorienting molecules. Examples include crystalline carbon tetrachloride at high temperatures, and methane. Then, the local correlation of the molecular orientations might be determined. This is NRC no. 28041.

Received 13 April; accepted 10 September 1987.

1. Peterson, S. W. & Levy, H. A. *Acta crystallogr.* **10**, 70–76 (1957).
2. Chamberlain, J. S., Moore, F. H. & Fletcher, N. H. in *Physics and Chemistry of Ice* (eds Whalley, E., Jones, S. J. & Gold, L. W.) 283–284 (Royal Society of Canada, Ottawa, 1973).
3. Kuhs, W. F. & Lehmann, M. S. *Nature* **294**, 432–434 (1981).
4. Kuhs, W. F. & Lehmann, M. S. *J. phys. Chem.* **87**, 4312–4313 (1983).
5. Kuhs, W. F. & Lehmann, M. S. *Water Sci. Rev.* **2** (1986).
6. Kern, C. W. & Karplus, M. in *Water a Comprehensive Treatise* Vol. 1 (ed. Franks, F.) 21 (Plenum, New York, 1972).
7. Whalley, E. *Molec. Phys.* **28**, 1105–1108 (1974).
8. Narten, A. H., Venkatesh, G. C. & Rice, S. A. *J. chem. Phys.* **64**, 1106–1121 (1976).
9. Floriano, M. A., Whalley, E., Svensson, E. C. & Sears, V. F. *Phys. Rev. Lett.* **57**, 3062–3064 (1986).
10. Mishima, O., Calvert, L. D. & Whalley, E. *Nature* **310**, 393–395 (1984).
11. Handa, Y. P., Mishima, O. & Whalley, E. *J. chem. Phys.* **84**, 2766–2770 (1986).
12. Leadbetter, A. J. *et al. J. chem. Phys.* **82**, 424–428 (1985).
13. Placzek, G. *Phys. Rev.* **86**, 377–388 (1952).
14. Jones, R. E. Report SC-M-69-335 (Sandia National Laboratories, Albuquerque, 1969).
15. Lorch, E. *J. Phys.* C**2**, 229–237 (1969).
16. Powles, J. G. *Molec. Phys.* **37**, 623–641 (1979); **42**, 757–765 (1981).
17. Klug, D. D. & Whalley, E. *J. chem. Phys.* **81**, 1220–1228 (1984).

Viscosity of sea-surface slicks

David J. Carlson

College of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

Theoretical models generally attribute altered wave properties in slicks on the ocean surface to elastic[1,2] or dissipative[3] effects within or attributable to monomolecular organic films. But the situation on the ocean surface is both dilute and complex: a variety of organic materials have been measured in surface samples[4] but their total abundance is lower than the surface abundances of pure compounds used in experiments to corroborate the monolayer models[5], and the organic materials appear, based on samples within the uppermost millimetre of the ocean, to be arranged in relatively thick microlayers[6,7] rather than in monolayers. Because observations of slicks correlated with microlayer enrichments of ultraviolet-absorbing materials[8], and slicks appeared viscous—even syrupy—when sampled, I suggested[8] that the presence of slicks involved viscosity effects in microlayers. Here I report fluorescence depolarization measurements of the bulk intrinsic viscosity of natural oceanic microlayers, showing that microlayer samples from slicked surfaces are indeed more viscous than samples from adjacent unslicked surfaces.

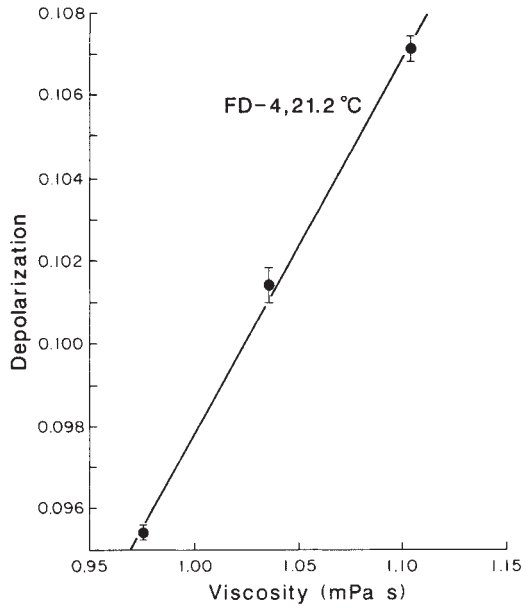


Fig. 1 Depolarization of FD-4 (FITC-dextran) as a function of viscosity of $MgSO_4$ solutions at 21.2 °C. Symbols show means and 95% confidence intervals for eight measurements of each solution. Equation describing the line is: $y = 0.0912x + 0.0066$, where y is depolarization and x is viscosity. In practice, the minimum possible value is the viscosity of pure water at 21.2 °C, 0.978 mPa s. Correlation coefficient for this single set of calibration standards was 0.980. This data set was selected to show precision and sensitivity in the range of normal sea-water values; linear calibrations were obtained for values to 1.30 mPa s.

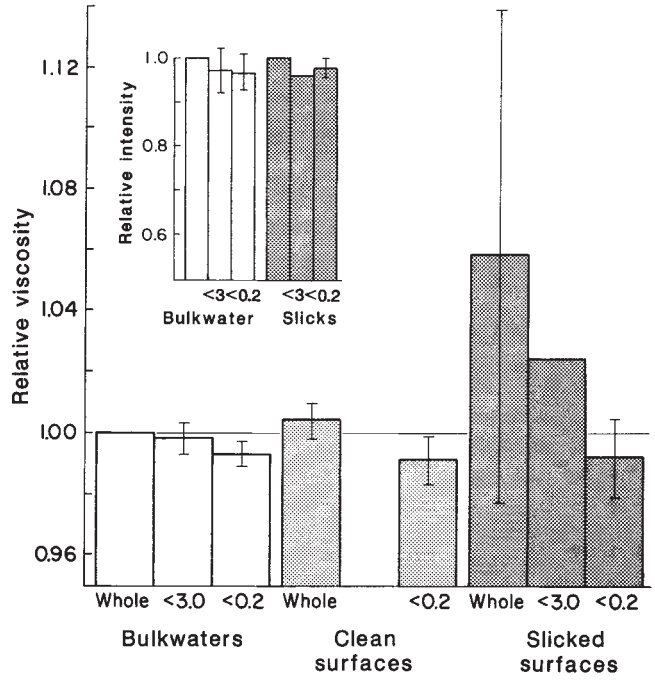


Fig. 2 Mean viscosities of whole (unfiltered), 3.0 μm filtered, and 0.2 μm filtered samples of bulkwaters and unslicked (clean) and slicked microlayers, relative to whole bulkwater. Sequence was: add dye to whole sample, measure, filter through 3.0-μm membrane, measure, filter through 0.2-μm membrane, measure. Error bars indicate 95% confidence intervals where $n > 2$. Inset shows unpolarized fluorescence intensity in the filtered bulkwater and slicked microlayer samples, relative to whole bulkwater and whole slicked microlayer samples respectively, with error bars as above.

Fluorescence depolarization techniques determine the rotational freedom of fluorescent molecules in solution: if the molecules are excited with polarized light, rotation of the excited molecule before emission will result in depolarization of the emitted light. For a fluorochrome of constant size and conformation in solution at constant temperature, the degree of depolarization is a direct indication of the solution viscosity and can be calibrated against solutions of known viscosity. The technique has been used to determine cytoplasmic[9] and membrane[10] viscosities and is uniquely suited to the accurate measurement of viscosity of seawater and microlayer samples containing surface-active organic materials and particulates[11]. The viscosity thus measured is an intrinsic bulk property of the microlayer sample, distinct from surface dilational or shear viscosities which are monolayer properties.

Microlayer samples were collected using glass plates[7] from clean and visibly slicked surfaces in coastal waters of the temperate Pacific and temperate and tropical Atlantic Oceans (Oregon, Maine and Florida) in August and September 1986. The samples were of representative natural slick regions, most less than a few metres in extent, present during the hour or two of sampling in each location. Subsurface waters were collected from 20 cm depth. All samples were contained unfiltered in silanized glass vials from collection until measurement. Viscosities of subsurface waters (bulkwater), unslicked (clean) microlayers and slicked microlayers were measured by fluorescence depolarization, using fluorescein isothiocyanate (FITC) bonded to a dextran (total molecular weight (MW), 4,100), added to 2 ml of bulkwater and microlayer samples to give a final concentration of 0.5 μM (39 nM as FITC). Fluorescence intensities were measured in a water-jacketed holder in a spectrofluorometer (excitation wavelength, 495 nm; emission wavelength, 518 nm), with polarizing filters in excitation and emission light paths. Four measurements were made, one for each possible combination of polarizer orientation: HH, HV, VH, VV, where H is horizontal, V is vertical, and the excitation polarizer is listed first. Depolarization was calculated as $(VV - G * VH)/(VV + G * VH)$, where $G = HV/HH$, a grating factor which corrects for polarization inherent in the emission detection system[12]. Depolarization of FITC-dextran was calibrated using $MgSO_4$ solutions of known viscosity (Fig. 1). Precision of the technique was generally ± 0.007 mPa s (95% confidence intervals for eight measurements, four each from two subsamples of each bulkwater or microlayer sample).

Viscosities of all slick microlayer samples from all Pacific and Atlantic locations were greater than viscosities of unslicked surfaces (Table 1); the latter were in most cases not significantly more viscous than subsurface waters. The largest absolute value (1.281 mPa s at 21.2 °C) was measured in a slick in Maine coastal waters. This value was 0.215 mPa s more viscous than that of the adjacent unslicked surface. Other slick viscosity enhancements were in the range 0.010 to 0.027 mPa s at 21.2 °C. Note that none of these samples was from as large, as dense, or as syrupy a slick as I have observed. The microlayer viscosities were persistent: those from Oregon waters did not change over three days and as many as five days elapsed while Atlantic samples were returned to Oregon for measurement. The viscosity of some slick samples decreased after filtration through 3.0-μm membranes; the viscosity of most slick samples decreased following filtration through 0.2-μm membranes (Fig. 2). It was not possible to distinguish between actual retention of viscous materials on the filters and disruption of viscous materials by the filtration process. Filtration reduced only depolarization, not fluorescence intensity, indicating that viscosities of the unfiltered samples were inherent properties rather than artefacts of dye association with filterable materials. Thus, microlayer samples from surfaces on which slicks were visible were more viscous than samples from unslicked surfaces.

The compositions of materials which contribute to the viscosity of slick microlayers are unknown. There was a correlation

Table 1 Viscosities (mean ± 95% confidence intervals) of clean and slicked surface microlayers and corresponding bulkwaters, measured by fluorescence depolarization techniques at 21.2 °C

Date	Location	Bulkwater (mPa s)	Clean surface (mPa s)	Slicked surface (mPa s)
14 August	Oregon	1.030 ± 0.011	1.029 ± 0.014	1.054 ± 0.024
25 August	Oregon	1.044 ± 0.007	1.039 ± 0.009	1.055 ± 0.007
	same	1.043 ± 0.007	1.043 ± 0.003	1.093 ± 0.016
17 Sept.	Maine	1.057 ± 0.003	1.061 ± 0.007	1.097 ± 0.007
	same	1.055 ± 0.007	1.066 ± 0.007	1.281 ± 0.007
18 Sept.	Maine	1.055 ± 0.007	1.059 ± 0.018	1.070 ± 0.016
	same	1.055 ± 0.008	1.064 ± 0.003	1.078 ± 0.007
25 Sept.	Florida	1.081 ± 0.017	1.082 ± 0.017	
	same	1.075 ± 0.005	1.067 ± 0.016	1.092 ± 0.005

Locations, dates, and original water temperatures: Oregon—Yaquina Bay, 14 and 25 August 1986, 13.5 °C; Maine—Damariscotta Estuary, 17 and 18 September 1986, 14.0 °C; Florida—off Tampa Bay, 25 September 1986, 29.5 °C. All salinities were ≥30‰. Microlayer samples, collected with a glass plate, had sample thicknesses of 40–60 μm (volume collected per dip/plate area). Bulkwaters were sampled at 20 cm depths. Samples were stored in silanized glass vials and returned to Oregon for depolarization measurements.

between microlayer viscosity and enrichment of UV-absorbing materials (microlayer absorbance/bulkwater absorbance)[8] but the absorbant materials passed 0.2-μm filters while the viscous materials generally did not. An appropriate model for composition and formation of viscous microlayers may include effects such as polymers have on solution properties. As little as 0.05 wt% of certain organic polymers can result in full gelation[13]; increased viscosity has been observed in non-gelling agarose solutions at solute volume fractions as small as 0.05% (ref. 14). Organic carbon in oceanic microlayers is in the range 0.001–0.01 wt% (taking microlayer total organic carbon as 5–50 mg l (ref. 15) and assuming the materials were 50% carbon). The disparity between amounts measured on the ocean surface and amounts required for laboratory gelation is only slight considering that: (1) microlayer viscosities sufficient to modify surface waves may be only a fraction of gel viscosities; (2) microlayer components such as proteins may have enhanced volumes due to orientation near the interface[16]; and (3) microlayers can also be enriched in major cations[17] and trace metals[18,19] whose effects may be to bridge and thereby extend the influence of individual organic molecules.

The presence of viscous microlayers seems to violate a central assumption of monolayer models, which is that bulk properties such as viscosity are constant right up to the actual interface[1]. Monolayer models also consider interfacial films as newtonian liquids, but the presence of filterable materials in the microlayer samples and their contribution to microlayer viscosity may indicate non-newtonian properties in the interfacial region[20]. It may be possible to assign microlayer viscosity to the surface as a component of surface shear or surface dilational viscosity[21], or it may be necessary to include explicit terms for vertical and horizontal dissipation between upper and lower boundaries of the microlayer[22], terms which are considered negligible or ignored in monolayer models[23]. The viscosities measured were intrinsic volume properties of the natural microlayer samples, evident without spreading or compressing the samples at an interface. As such, the microlayer viscosities are distinct from vicinal water theories[24] which require compressed monolayer films to induce near-surface viscosity.

Models for gas[25], momentum[26] and heat exchange[27] across the air–sea interface all invoke viscous microlayers, as so-called stagnant films or thermal or viscous sublayers, of dimensions similar to depths of microlayers sampled here. It is evident from the data presented here that, in slicks at least, the viscosity of those sublayers may be considerably greater than subsurface viscosity. The importance of such viscosity enhancements on air–sea exchange depends on the extent of surface slick conditions. The extent of viscous slick conditions will also influence emissivity of the ocean surface: relaxation times of water itself[28] and of solutes such as proteins[29] increase directly with viscosity. Increased relaxation times in viscous microlayers will decrease dielectric constants and brightness temperatures[30], effects previously attributed to long-distance ordering induced by monolayer films[31]. It is clear that discussion and investigation of interfacial properties and processes must be expanded to include consideration of viscous microlayers.

Thanks to L. Mayer, J. Paul and M. DeFlaun for help in sampling, to L. Morrill and J. Brophy for the analyses and to John C. Scott for a very thorough and constructive review. This work supported by the Office of Naval Research.

Received 9 March; accepted 9 September 1987.

1. Lucassen-Reynders, E. H. & Lucassen, J. *Adv. Colloid Interface Sci.* **2**, 347–395 (1969).
2. Ermakov, S. A. *et al. Dynam. Atmos. Oceans* **10**, 31–50 (1986).
3. Pogorzelski, S., Linde, B. & Sliwinski, A. *Acoust. Lett.* **8**, 5–9 (1984).
4. Williams, P. M. *et al. Mar. Chem.* **19**, 17–98 (1986).
5. Hunter, K. A. & Liss, P. S. in *Marine Organic Chemistry* (eds Duursma, E. K. & Dawson, R.) 259–298 (Elsevier, Amsterdam, 1981).
6. Daumas, R. A., Laborde, P. L., Marty, J. C. & Saliot, A. *Limnol. Oceanogr.* **21**, 319–326 (1976).
7. Carlson, D. J. *Mar. Chem.* **11**, 189–208 (1982).
8. Carlson, D. J. *Nature* **296**, 426–429 (1982).
9. Rotman, A. & Heldman, J. *Biochemistry* **20**, 5995–5999 (1981).
10. Fuchs, P., Parola, A., Robbins, P. W. & Booth, E. R. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3351–3354 (1975).
11. Carlson, D. J., Morrill, L. E. & Brophy, J. E. *Limnol. Oceanogr.* (in the press).
12. Chen, R. F. & Bowman, R. L. *Science* **147**, 729–732 (1965).
13. Sinclair, M., Lim, K. C. & Heeger, A. J. *Phys. Rev. Lett.* **51**, 1768–1771 (1983).
14. Letherby, M. R. & Young, D. A. *J. chem. Soc. Faraday Trans.* **77**, 1953–1966 (1981).
15. Carlson, D. J. *Limnol. Oceanogr.* **28**, 415–431 (1983).
16. MacRitchie, F. *Adv. Colloid Interface Sci.* **25**, 341–385 (1986).
17. Pattenden, N. J., Cambray, R. S. & Playford, K. *Geochim. cosmochim. Acta* **45**, 93–100 (1981).
18. Barker, D. R. & Zeitlin, H. *J. geophys. Res.* **77**, 5076–5086 (1972).
19. Hardy, J. T., Apts, C. W., Crecelius, E. A. & Bloom, N. S. *Est. Coastal Shelf Sci.* **20**, 299–312 (1985).
20. Jenkinson, I. R. *Nature* **323**, 435–437 (1986).
21. Ting, L., Wasan, D. T., Miyano, K. & Xu, S.-Q. *J. Colloid Interface Sci.* **102**, 248–259 (1984).
22. Mordashev, V. I. & Cherkesov, L. V. *Izvestia Atmos. Ocean Phys.* **22**, 306–310 (1986).
23. Goodrich, F. C. *J. phys. Chem.* **66**, 1858–1863 (1962).
24. Huhnerfuss, H. & Alpers, W. *J. phys. Chem.* **87**, 5251–5258 (1983).
25. Liss, P. S. *Deep-Sea Res.* **20**, 221–238 (1973).
26. Wu, J. *J. phys. Oceanogr.* **14**, 138–144 (1984).
27. Paulson, C. A. & Simpson, J. J. *J. geophys. Res.* **86**, 11044–11054 (1981).
28. Collie, C. H., Hasted, J. B. & Ritson, D. M. *Proc. Phys. Soc.* **60**, 145–160 (1948).
29. Hasted, J. B. *Aqueous Dielectrics* (Chapman & Hall, London, 1973).
30. Maul, G. A. *Introduction to Satellite Oceanography* (Nijhoff, Dordrecht, Netherlands, 1985).
31. Alpers, W., Blume, H.-J. C., Garrett, W. D. & Huhnerfuss, H. *Int. J. Remote Sensing* **3**, 457–474 (1982).

Chernobyl radionuclides in a Black Sea sediment trap

K. O. Buesseler,*, H. D. Livingston*, S. Honjo* B. J. Hay*, S. J. Manganini*, E. Degens†, V. Ittekkot†, E. Izdar‡ & T. Konuk‡

* Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA
† Geologisch Paläontologisches Institut, Universität Hamburg, D-2000 Hamburg 13, FRG
‡ Dokuz Eylül Universitesi, Deniz Bilimleri ve Teknolojisi Enstitüsü, Izmir, Turkey

The Chernobyl nuclear power station accident[1–3] released large quantities of vaporized radionuclides, and, to a lesser extent, mechanically released small (<1–10 μm) aerosol particles[2,4]. The total release of radioactivity is estimated to be of the order of $1–2 \times 10^{18}$ Bq ($3–5 \times 10^{7}$ Ci) not allowing for releases of the xenon and krypton gases[2]. The ^{137}Cs releases of 3.8×10^{16} Bq from Chernobyl can be compared to 1.3×10^{18} Bq ^{137}Cs released due to atmospheric nuclear weapons testing[1]. Chernobyl-derived radionuclides can be used as transient tracers to study physical and biogeochemical processes. Initial measurements of fallout Chernobyl radionuclides from a time-series sediment trap at 1,071 m

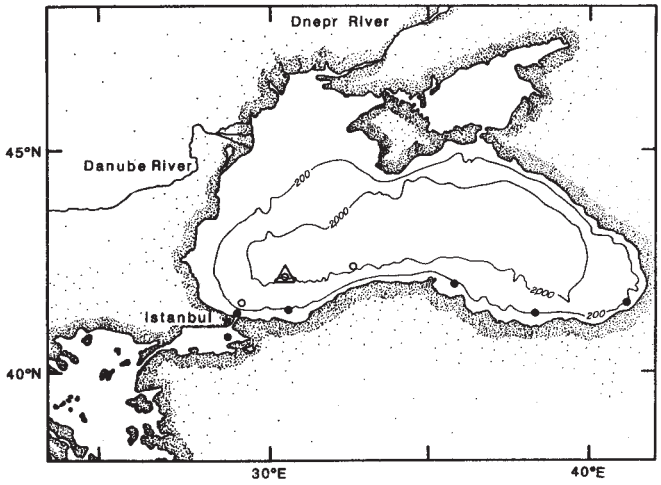


Fig. 1 Location map for Black Sea samples. The triangle marks the sediment-trap site. The open and filled circles represent the location of surface-water samples for particulate and dissolved radionuclides. The open circles are deep stations (>200 m), while the filled circles are shelf stations (<200 m). Depth contours are in m.

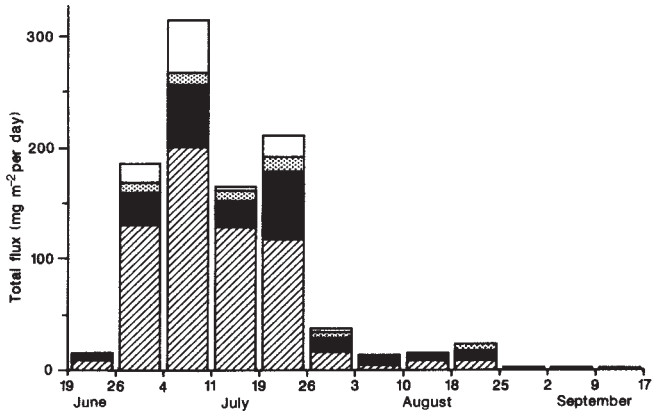


Fig. 2 Total flux (in mg m^{-2} per day) of settling particles found in each sampling cup between 19 June and 17 September 1986. Each flux bar is divided into the percentage of lithogenic, carbonate and biogenic silica found in the sample. ▨, Carbonate; ■, organic carbon; ▩, biogenic silica; □, lithogenics.

during June–September 1986 in the southern Black Sea are presented. The specific activities of ^{137}Cs, ^{144}Ce and ^{106}Ru in the trap samples (0.5–2, 4–12 and 6–13 Bq g^{-1}) are independent of the particle flux while their relative activities reflect their rates of scavenging in the order Ce > Ru > Cs.

The Black Sea is the closest salt-water body to the Chernobyl reactor site and is also connected to the accident site by the Dnepr River. Fallout radioactivity was likely to have been deposited to the Black Sea predominantly during the later stages of the Chernobyl accident (1–6 May)[5]. Chernobyl fallout in southern Europe[6] and the Black Sea region is more enriched in the refractory radionuclides such as ^{144}Ce and ^{106}Ru, relative to the more volatile ^{131}I and ^{137}Cs (which were more abundant in the fallout releases before May[4,6]). Elevated ^{134}Cs and ^{137}Cs levels in June in the surface waters of the southern Black Sea[7] confirm the existence of widespread direct atmospheric deposition of fallout from the Chernobyl accident to this basin. With time, runoff from the Danube and Dnepr Rivers will provide an additional indirect source of Chernobyl radioactivity for this basin.

The Black Sea is a unique water body in that the surface oxic waters (which extended down to 120 m at the trap site in September) are underlain by a large basin of permanently anoxic deep water. These two layers are separated by a strong density gradient between the brackish surface waters and the more saline deep waters.

A time-series sediment trap[8] with a 1.2-m^2 aperture was deployed from the RV *K. Piri Reis* at 41°51.22′ N, 30°21.15′ E (Fig. 1) at a depth of 1,071 m along a bottom-tethered mooring array (anchor depth 2,100 m). The trap collected settling particles over 12 equal intervals of 7.5 days each, from 20 June to 18 September 1986. The sample-receiving cups were poisoned with formaldehyde to inhibit biological activity. The trap was immediately redeployed for further monitoring of the radionuclide flux. Upon recovery, the sample-collection cups were sealed and transported to Woods Hole where they were kept under refrigeration. The following Chernobyl radionuclides were detected by non-destructive gamma spectrometry: ^{134}Cs, ^{137}Cs, ^{141}Ce, ^{144}Ce, ^{103}Ru, ^{106}Ru, ^{95}Nb, and ^{110m}Ag (with 6–18 h counting intervals in the energy range of 134–884 keV). The detector-counting rates were calibrated with a ^{152m}Eu solution standard. All activities are corrected for decay to 1 May 1986. The presence of short-lived ^{141}Ce, ^{103}Ru, ^{95}Nb and ^{134}Cs/^{137}Cs activity ratio of 0.5 (characteristic of Chernobyl fallout) in the trap material allowed us to positively identify the radioactivity flux as resulting from the recent Chernobyl releases.

The total sediment-trap flux increased from 19 mg m^{-2} per day in June to a maximum of 360 mg m^{-2} per day during the week of 4–11 July 1986. The flux remained high throughout July, decreased in August, and dropped off to <3 mg m^{-2} per day in September. This seasonal peak in particle flux is consistent with previous sediment-trap studies at a nearby site[9] and is due to the high productivity of the coccolithophorid species, *Emiliania huxleyi*. The trap material in the bloom period consists of >60% calcium carbonate, and is composed primarily of coccoliths (Fig. 2). In contrast to other periods of the year[9], the flux of biogenic silica and lithogenic particles is relatively low during the sampling period (Fig. 2).

The radiochemical data (Table 1 and Figs 3 and 4) are given for ^{134}Cs, ^{137}Cs, ^{144}Ce and ^{106}Ru as these radionuclides were readily detected by the gamma-spectrometry technique, and their half-lives are long enough to allow the study of their environmental geochemistry over a period of a few to several tens of years ($t_{1/2}$ = 2.07, 30.17, 0.78 and 1.02 yr, respectively). To facilitate reference to other studies, we report here our mean ^{103}Ru/^{106}Ru and ^{141}Ce/^{144}Ce ratios to be 4.80 ± 0.07 and 1.44 ± 0.06, respectively. The specific activities of ^{137}Cs, ^{144}Ce and ^{106}Ru in the trap samples contrast sharply with the total flux pattern and are not identical for each isotope (Fig. 3 and Table 1). The specific activity of ^{137}Cs decreases slightly from ≈2 Bq g^{-1} (dry weight) to <1 Bq g^{-1} between June and August. The ^{144}Ce specific activity in mid-June is 6.2 Bq g^{-1}, increasing systematically to a peak of almost 12 Bq g^{-1} at the end of July, and then rapidly decreasing to the 4–5 Bq g^{-1} level. The ^{106}Ru specific activity is roughly constant at 7 Bq g^{-1} with a single point activity maximum in late August. The mechanism responsible for these trends is not reflected in the patterns shown in total mass and other sedimentary components (compare Figs 2 and 3). The lack of correlation between trap flux and specific activity is consistent with other studies of particle reactive elements[10,11] and reflects, at least partially, a constant exchange equilibrium between the dissolved and particulate phases. This equilibrium is between the dissolved phase, the suspended particulate (small) and settling particulate (large) phases. Radionuclides exchange reversibly between the dissolved and suspended phases, and the suspended and settling particulate phases exchange by aggregration/disaggregation. Hence, the radionuclide flux is proportional to the settling particle flux. The total radionuclide fluxes are driven by the much larger seasonal variation in primary productivity, which is reflected in the large flux of biogenic components during the first half of the sampling period (Table 1).

The ^{134}Cs/^{137}Cs, ^{144}Ce/^{137}Cs and ^{106}Ru/^{137}Cs ratio data can

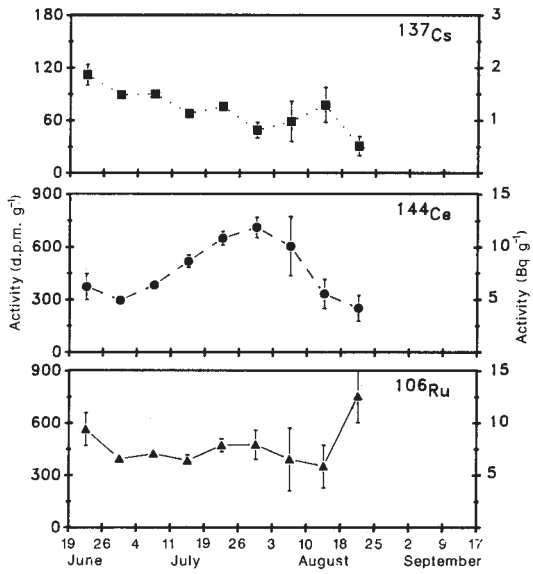


Fig. 3 Sediment trap ^{137}C, ^{144}Ce and ^{106}Ru specific activities against time for the sediment-trap material collected between 19 June and 17 September 1986 (data in Table 1). The dates on the *x*-axis are the opening and closing dates for the sediment-trap sample cups. The activity scales are given as d.p.m. g^{-1} on the left vertical axis and as Bq g^{-1} on the right vertical axis (60 d.p.m. = 1 Bq). The error bars indicate the analytical error given in Table 1 and are drawn if larger than the symbol.

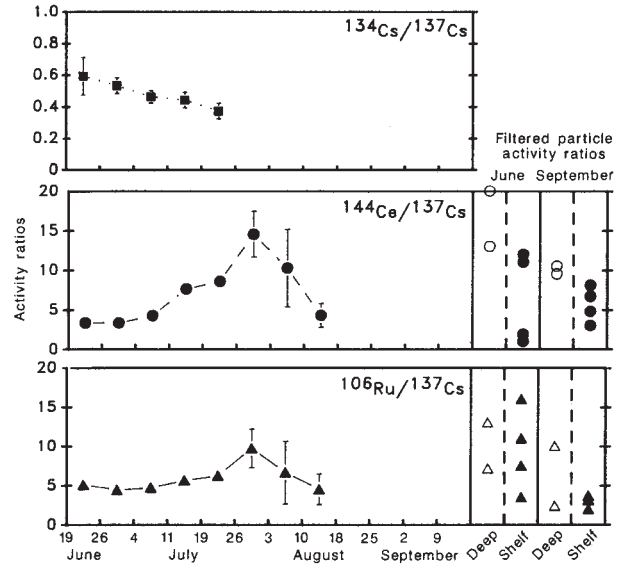


Fig. 4 The sediment-trap $^{134}Cs/^{137}Cs$, $^{144}Ce/^{137}Cs$ and $^{106}Ru/^{137}Ca$ activity ratios against time. The ratios shown are from the first 5–8 samples in the time series only, as the error on the activity ratios (drawn as an error bar around the data points) becomes unacceptably large in the later samples which had the lowest sample masses (see Table 1). For ^{144}Ce and ^{106}Ru, where the radionuclide activity on the filtered particles (>0.45 μm) is substantial, their filtered particle activity ratios are also shown for comparison on an extension to the right-hand side of the graphs. In each case, the ratio data from the filtered particles are given from the June (left) and September (right) cruises and from surface waters collected at deep (open symbols) and shelf (filled symbols) stations (see Fig. 1 for station locations).

be used to examine the relative scavenging chemistries of Cs, Ce and Ru (Fig. 4). There are two sources of fallout Cs in the Black Sea: Chernobyl fallout with its characteristic $^{134}Cs/^{137}Cs$ ratio of ~0.5, and pre-existing atmospheric nuclear weapons fallout with a $^{134}Cs/^{137}$ Cs ratio of zero (due to ^{134}Cs decay since the major bomb inputs in the 1960s). The $^{134}Cs/^{137}Cs$ ratio in the trap material decreases with time from an initial value of ~0.5 to <0.440. This decrease in the $^{134}Cs/^{137}Cs$ ratio and the total ^{137}Cs activity with time indicates that the percentage of Chernobyl ^{137}Cs, relative to global fallout ^{137}Cs is decreasing in the trap material during this collection period.

The $^{144}Ce/^{137}Cs$ ratio in the sediment-trap samples (Fig. 4) peaks around the end of July, reflecting the strong trend in Ce specific activity discussed previously. In Fig. 4, the $^{144}Ce/^{137}$ Cs ratio of the trap material is compared to the $^{144}Ce/^{137}Cs$ ratio we have found in filtered particles (>0.45 μm) collected in the surface waters of the Black Sea. If we consider the fine particles which are collected in the overlying surface waters as being the source of the larger aggregates found in the trap, or at least as being similar in the particle-to-solution partitioning, then two features should be noted. First, the range of $^{144}Ce/^{137}Cs$ ratios found in the fine particles and sediment-trap samples is similar. Second, the surface-water particulate $^{144}Ce/^{137}Cs$ ratios vary, depending upon their location, such that the $^{144}Ce/^{137}Cs$ ratios of the shelf stations are lower than those at the deeper stations. This is apparently due to increased Ce removal in the more

Table 1 Black Sea sediment trap at 1,071 m, 41°51.22′ N; 30°21.15′ E

Trap collection dates (1986)	Total flux mg m^{-2} day^{-1}	Specific activity (Bq g^{-1})* ^{137}Cs	^{144}Ce	^{106}Ru	Radionuclide flux (Bq m^{-2} per day) ^{137}Cs	^{144}Ce	^{106}Ru
19–26 June	18.9	1.9±0.2	6.2±1.2	9.4±1.6	0.04	0.12	0.18
26 June–5 July	231.1	1.5±0.08	5.0±0.3	6.6±0.5	0.34	1.15	1.52
4–11 July	356.9	1.5±0.08	6.4±0.3	7.1±0.5	0.54	2.28	2.52
11–19 July	184.7	1.1±0.08	8.7±0.6	6.4±0.6	0.21	1.60	1.18
19–26 July	230.9	1.3±0.08	10.9±0.6	7.9±0.6	0.29	2.51	1.82
26 July–3 August	45.8	0.8±0.2	11.9±0.9	7.9±1.4	0.04	0.54	0.36
3–10 August	15.2	1.0±0.4	10.1±2.8	6.5±3.0	0.01	0.15	0.10
10–18 August	16.2	1.3±0.3	5.6±1.4	5.9±2.1	0.02	0.09	0.09
18–25 August	30.3	0.5±0.2	4.2±1.2	12.6±2.5	0.02	0.13	0.38
25 August–2 September	2.7	1.9±1.4	8.2±5.6	9.0±10.3	0.01	0.02	0.02
2–9 September	2	BD	BD	BD	—	—	—
9–17 September	1.4	BD	BD	BD	—	—	—
			Annual flux (Bq m^{-2} yr^{-1}) =		46.0	261.1	248.7

* Specific activities as Bq g^{-1} dry weight, with one standard deviation counting error. All activities are decay corrected to 1 May 1986.
BD, below detection (due to low sample mass).
Estimated inventories of Chernobyl ^{137}Cs, ^{144}Ce and ^{106}Ru (data taken from our Black Sea Chernobyl studies[13]): Mean ^{137}Cs, ^{144}Ce, and ^{106}Ru activities in surface water are 170, 8, and 10 Bq m^{-3}, respectively. Penetration of ^{137}Cs into water column is 35 m (assume the same for ^{144}Ce and ^{106}Ru). Total inventories of ^{137}Cs, ^{144}Ce, and ^{106}Ru are 6.0×10^3, 2.8×10^2, and 3.5×10^2 Bq, respectively. Removal rate: Inventory/annual flux, ^{137}Cs = 130 years, ^{144}Ce = 1.1 years, ^{106}Ru = 1.4 years.

coastal sites (see Fig. 1 for location of shelf and deep-water stations). As the settling time of the sinking coccoliths is of the order of two weeks to reach 1,071 m^9, particles collected in the trap will not have been formed in the surface waters directly above the trap site due to rapid circulation of Black Sea surface waters (0.1-0.3 m s^{-1}, ref. 12). Therefore, an increase in the $^{144}Ce/^{137}Cs$ ratio could reflect a shift in the source of the trap particles from shelf to deeper sites. If, in general, the $^{144}Ce/^{137}Cs$ ratio is lowered in the surface waters due to continued Ce scavenging, then a longer trap record should show an overall decrease in the $^{144}Ce/^{137}Cs$ ratio with time.

The $^{106}Ru/^{137}Cs$ ratios in the sediment-trap material do not vary substantially and are within the range of the fine particle $^{106}Ru/^{137}Cs$ ratios (Fig. 4*b*). While Ru is more particle reactive than Cs, it is less particle reactive than Ce in the surface waters (the particulate/soluble ratios for ^{137}Cs, ^{106}Ru and ^{144}Ce are ~0.1%, 4% and 11%, respectively, at this trap site in September)[13,14] Consistent with this trend in particle reactivity (Cs < Ru < Ce), we have found that the relative enrichment of Ru on the trap material is less than that of Ce when compared to their total activities in the surface waters[13,14].

The trap-radionuclide fluxes can be compared to estimated water column Chernobyl-fallout inventories to examine the relative removal rates of each radionuclide. The removal time of Chernobyl ^{137}Cs (130 y) is much longer than that of ^{106}Ru and ^{144}Ce (1.4 and 1.1 yr, respectively, see Table 1). The latter could be an upper limit if significant amounts of Ce or Ru are soluble in the deep anoxic waters. The total ^{137}Cs flux measured over the first three months following Chernobyl is therefore rather small relative to the total ^{137}Cs inventory. The more particle-reactive Ru and Ce are being much more rapidly removed from the surface Black Sea, as expected by their affinity for the particulate phases. Relative particle-reactivity trends for Ce and Ru may be indicated better from the partition between particulate/soluble phases than from their removal-time estimates. The former approach does not require the added assumption of the association of observed particle fluxes with water-column inventories in September.

The initial sediment-trap results indicate that there was rapid removal of a considerable fraction of the Chernobyl radionuclides to the deep Black Sea. The more soluble isotope ^{137}Cs is less efficiently removed from the surface waters at this deep trap site than are the more particle-reactive isotopes, ^{144}Ce and ^{106}Ru. Future radiochemical analyses of ^{90}Sr (a highly soluble tracer) and the Pu, Am and Cm isotopes (particle-reactive tracers) in the trap samples will provide an additional range of contrasting geochemistries for comparison.

Coupled with water-column and sediment studies, the time-series sediment-trap record will allow us to determine the geochemical fate of recently introduced Chernobyl radionuclides in the Black Sea. Also, this work will have further important implications for understanding the geochemistries of a wide variety of naturally occurring chemical analogues in this unique oceanographic setting.

We thank our scientific colleagues for their cooperation, and the US National Science Foundation, the US Office of Naval Research, the Coastal Research Center at Woods Hole, the German Minister for Research and Technology, and the Scientific and Technical Research Council of Turkey for financial assistance. The officers and crew of the *K. Piri Reis* were invaluable in their assistance in deploying the sediment-trap mooring, and in sampling the Black Sea waters on two cruises in 1986. We thank our co-workers W. R. Clarke and S. A. Casso for their assistance in sample collection and radiochemistry. This is contribution number 6473 from the Woods Hole Oceanographic Institution.

Received 4 May; accepted 13 August 1987.

1. Levi, B. G. *Physics Today* December, 17-20 (1986).
2. *Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident.* Safety Series No. 75-INSAG-1 (IAEA, Vienna, 1986).
3. Hohenemser, C. *et al. Environment* **28**, 6-13, 30-43 (1986).
4. Devell, L. *et al. Nature* **321**, 192-193 (1986).
5. Alexandropoulos, N. G. *et al. Nature* **322**, 779 (1986).
6. Krey, P. W. *US Department of Energy Report EML-460* 155-213 (US Department of Energy, Environmental Measurements Laboratory, New York, 1986).
7. Livingston, H. D. *et al. US Department of Energy Rep. EML-460*, 214-223 (US Department of Energy, Environmental Measurements Laboratory, New York, 1986).
8. Honjo, S. & Doherty, K. *Deep Sea Res.* (in the press).
9. Honjo, S. *et al. Particle Flux in the Ocean, Mitteil. Geol. Pal. Inst.* Vol. 62 (eds Degens, E. T., Izdar, E. & Honjo, S.) (University of Hamburg, in the press).
10. Bacon, M. P., Huh, C.-A., Fleer, A. P. & Deuser, W. G. *Deep Sea Res.* **32**, 273-286 (1985).
11. Fowler, S. W. & Knauer, G. A. *Progr. Oceanog.* **16**, 147-194 (1986).
12. Zentovitch, V. T. in *The Encyclopedia of Oceanography, Encyclopedia of Earth Science Series* Vol. 1 (ed. Fairbridge, R. W.) 145-150 (Halsted, New York, 1966).
13. Livingston, H. D. & Buesseler, K. O. *Eos* **67**, 1070 (1986).
14. Livingston, H. D., Buesseler, K. O., Izdar, E. & Konuk, T. *Proc. Int. Symp. on Radioactivity and Oceanography* (Elsevier, in the press).

Chernobyl nuclide record from a North Sea sediment trap

S. Kempe* & H. Nies†

* Geologisch-Paläontologisches Institut, University of Hamburg, Bundesstrasse 55, D-2000 Hamburg 13, FRG
† Deutsches Hydrographisches Institut, Bernhard-Nocht-Strasse 78, D-2000 Hamburg 4, FRG

Nuclides liberated by explosion and subsequent fire at Chernobyl No. 4 reactor on 26 April 1986, travelled to Western Europe with lower tropospheric air masses[1]. They reached the northern Alps[2] and Paris[3] on 30 April, southern Great Britain[4] on 2 May, the southern North Sea[5] on 3 May, and the northern North Sea[1] on 3 May and again on 8 May. Levels of deposited activity varied by a factor of 30 or more over distances < 100 km because of variability of rainfall[2,6]. A sediment trap deployed 222-m-deep in the North Sea off Bergen recorded the onset and magnitude of the deposition of Chernobyl nuclides. The trap collected 13 samples between 24 April and 21 September 1986. The flux of nuclides adsorbed to particles sinking from surface waters to sediments started less than ten days after contaminated air reached the site. Maximum specific activity occurred on 16–27 May for ^{137}Cs, ^{134}Cs, ^{106}Ru and ^{103}Ru, and on 8–20 June for ^{144}Ce, ^{95}Nb and ^{95}Zr. The highest activity was found for ^{103}Ru. The highest total specific activity of these nuclides in depositing sediments reached 670,000 Bq kg^{-1}, and the highest total activity flux for one day amounted to 50 Bq m^{-2}.

A sediment trap recorded the onset and magnitude of the nuclide flux from the North Sea surface to sediments; it was stationed 30 nautical miles (1.852 km) off Bergen at the western side of the Norwegian Channel (60° 29.6′ N, 3° 50.4′ E, water depth 318 m). The Mark-VI Honjo sediment trap[7] was moored at a water depth of 222 m, well below the summer thermocline. Its 0.509-m^2 opening directed the settling particles into cups, which were automatically advanced at preprogrammed dates (interval 11.62 days during deployment NS III, summer 1986). The cups were filled with saturated NaCl- solution and poisoned with $HgCl_2$. The first cup was opened on 24 April and the last was closed on 20 September 1986 (Table 1). The trap and samples were recovered by the German RV *Valdivia*, and the samples were transported undisturbed and cooled.

The samples were investigated for Chernobyl nuclides on 20-27 October by gamma-spectroscopy at the DHI. Original sample cups were placed directly on the Ge(Li)-detectors. The energy-dependent efficiency calibration was done with a nuclide mixture of the Physikalisch-Technische-Bundesanstalt, which contained known activities of ^{7}Be, ^{54}Mn, ^{57}Co, ^{65}Zn, ^{88}Y, ^{133}Ba, ^{137}Cs and ^{139}Ce. It was assumed that all of the activity was adsorbed to settled sediment covering the bottom of the cups and had not dissolved to any extent in the overlying NaCl-solution. This assumption was later tested with filtered sediment. Both measurements delivered comparable (decay-corrected) results, suggesting that, in fact, all of the measured activity was bound to particles. The counting time for samples N3 to N10

Table 1 North Sea sediment trap Norwegian Channel, deployment NSIII, summer 1986: samples dates, amounts, fluxes and composition of trapped material

1 Sample cup	2 Date open	3 Mass (mg)	4 Corrected mass (mg)	5 Flux ($mg\ m^{-2}\ d^{-1}$)	6 C_{org} (mg)	7 $CaCO_3$ (mg)	8 Terrigenous + opal (mg)
1	24.04.86	564.0	686.6	127.0	72.37	20.03	485.65
2	04.05.86	1129.0	1374.4	232.4	148.44	111.10	892.21
3	16.05.86	124.0	151.0	25.5	23.69	9.06	82.72
4	27.05.86	308.0	375.0	63.4	48.79	89.38	163.65
5	08.06.86	54.0	65.7	11.1	27.17	—	0.00
6	20.06.86	354.5	431.6	73.0	58.87	225.15	59.28
7	01.07.86	1024.5	1247.2	210.9	95.29	852.25	156.73
8	13.07.86	11.0	13.4	2.3	3.84	—	3.80
9	25.07.86	81.5	99.2	16.8	46.27	—	0.00
10	05.08.86	104.0	126.6	21.4	54.83	40.72	0.00
11	17.08.86	19.0	23.1	3.9			
12	28.08.86	21.0	25.6	4.3		not enough	
13	09.09.86	13.0	15.8	2.7		material avail-	
close	20.09.86					able for analysis	
14	—	827.7	—	—			

Table 2 Specific activity of sediment trap material, activities are decay corrected to 1 May 1986, errors are given in ±1 s.d. in percentage of mean

Cup no.	^{137}Cs ($Bq\ g^{-1}$)	Error	^{134}Cs ($Bq\ g^{-1}$)	Error	^{106}Ru ($Bq\ g^{-1}$)	Error	^{103}Ru ($Bq\ g^{-1}$)	Error
1	—	—	—	—	—	—	—	—
2	4.7	2.1	2.5	2.3	14.6	5.5	176	3.2
3	13.6	5	7.4	9	54.9	14	561	6
4	6.1	7	3.3	10	25.3	10	233	12
5	6.4	10	3.4	16	33.5	35	307	18
6	2.4	15	1.8	18	10.7	34	97.1	22
7	1.6	5.4	0.9	7.5	4.8	11	60.1	10
8	3.6	64	—	—	—	—	—	—
9	1.1	54	—	—	—	—	72.6	57
10	2.5	29	—	—	—	—	—	—
11	3.5	28	3.4	60	—	—	227	48
12	6.0	43	5.4	43	—	—	194	54
13	11.2	31	5.6	67	—	—	—	—

Cup no.	^{144}Ce ($Bq\ g^{-1}$)	Error	^{141}Ce ($Bq\ g^{-1}$)	Error	^{95}Nb* ($Bq\ g^{-1}$)	^{95}Zr ($Bq\ g^{-1}$)	Error	^{95}Zr† ($Bq\ g^{-1}$)	Error
1	—	—	—	—	—	—	—	—	—
2	2.1	14	14.1	48	5.0	5.4	18	10.4	5.6
3	12.9	35	—	—	16.1	18.6	40	34.7	15
4	3.3	73	—	—	—	—	—	8.4	28
5	31.2	30	—	—	37.1	36.6	36	74.0	14
6	—	—	—	—	—	—	—	—	—
7‡	1.1	43			—	—	—	1.0	28

* Difference of measured ^{95}Zr and ^{95}Zr calculated from transient equilibrium between ^{95}Zr and ^{95}Nb.
† Calculated from transient equilibrium with ^{95}Nb.
‡Cups 8–13 below limit of detection.

and N13 was 1,000 min, and for N1, N2, N11 and N12 it was 3,600 min. Table 2 gives results of measurements as specific activities ($Bq\ g^{-1}$). Activities are corrected for decay to 1 May 1986, the middle of the release period of nuclides from Chernobyl (26 April–5 May)[8]. ^{95}Zr disintegrates to ^{95}Nb representing a transient equilibrium; if one computes the activity of ^{95}Zr as of 1 May based on the activity of ^{95}Nb, it is roughly a factor of two higher than if measured directly. This reveals that these two isotopes have been formed in equivalent amounts during the reactor accident. With this method it is possible to calculate the activity in those samples in which no ^{95}Zr was detected. The statistical counting error (coefficient of variation) decreases with increasing gamma-emission probability and with increasing activity of the different nuclides in the sample. Because of a decrease in particle flux towards the end of the summer, counting errors increase in the latter cups (Table 2).

After activity measurement, samples were sieved. Fractions >1 mm contained only exuvia of crustaceans. Fractions <1 mm were split wet into ten aliquots; depending on total amount, two or four of the aliquots were filtered, washed and dried at 40 °C and used to determine total mass and concentrations of organic and inorganic carbon (Table 1). C_{inorg} results have been recalculated for $CaCO_3$, and C_{org} data were used to estimate total organic matter (assumed to be CH_2O). The remainder of total amount minus the sum of $CaCO_3$ and organic matter gives an estimate of terrigenous matter plus biogenic opal (Table 1).

The relatively shallow deployment of the trap poses the problem of colonization of trap surfaces by marine organisms. These produce a sticky coating retaining some of the intercepted particles. Upon recovery, 840 mg of material was removed from the inner cone surface (sample N14) with an unknown amount lost during recovery; thus fluxes recorded by sediment traps rep-

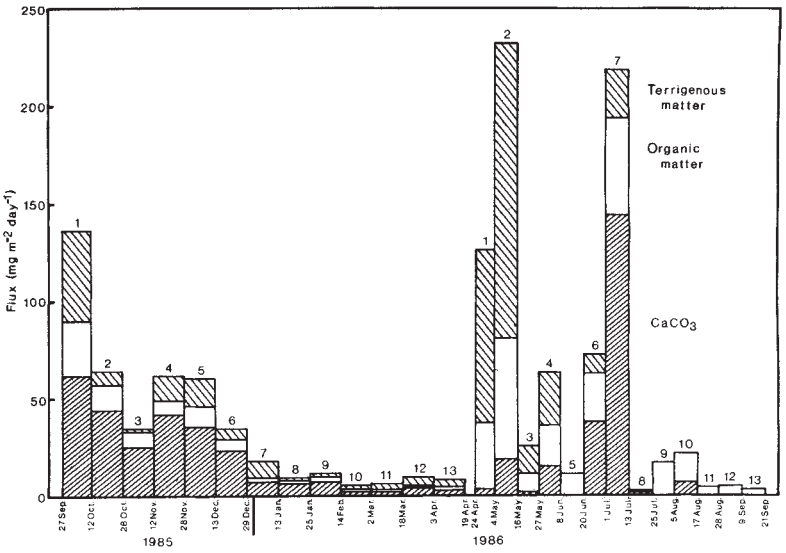


Fig. 1 Sediment flux (grainsize <1 mm) at site Norwegian Channel from deployments NSII and NSIII, September 1985–September 1986 and its composition as to $CaCO_3$, organic matter and terrigenous matter (including biogenic opal).

resent a lower limit of actual fluxes. The cone sample was added to the amount of sediment in the individual cups (Table 1, column 3) in proportion to the respective amount (column 4) and total yields were recalculated for daily flux (mg m^{-2} d^{-1}) (column 5).

The total amount of material that settled during the year September 1985–September 1986 was 16.7 g m^{-2} or 46 mg m^{-2} d^{-1} (Fig. 1). The highest rates were recorded on 4–16 May 1986, and the lowest on 13–25 July 1986, amounting to 232.4 and 1.15 mg m^{-2} d^{-1}, respectively. Thus fluxes are highly variable (by a factor of 200). During winter, rates decreased from 140 mg m^{-2} d^{-1} in late September to about 5 mg m^{-2} d^{-1} in late winter and early spring. Because of extremely low winter temperatures in 1985–86, the spring plankton bloom was late throughout the North Sea and did not occur until late April–early May as evident from the high flux values in the first two samples in Table 1 (Fig. 1). The flux of terrigenous material is highest during spring plankton bloom, whereas $CaCO_3$ flux is highest in summer and autumn. The high percentage of biogenic material in the Norwegian Channel trap is connected to regional currents. The trap was moored below a current which brings water of the open North Atlantic[9] south into the North Sea.

The highest activities of the Chernobyl nuclides in the North Sea were shown by ^{137}Cs, ^{134}Cs, ^{131}I, ^{132}Te/^{132}I, ^{103}Ru and ^{140}Ba/^{140}La in early May 1986[5]. In the trap material eight gamma-emitting Chernobyl nuclides could still be detected in October 1986: ^{137}Cs, ^{134}Cs, ^{106}Ru, ^{103}Ru, ^{144}Ce, ^{141}Ce, ^{95}Nb and ^{95}Zr (Table 2, Fig. 2*a*). These nuclides have half-lives of 30 yr, 2.1 yr, 373 d, 39.35 d, 284 d, 32.5 d, 35.1 d and 64 d, respectively. The averaged ^{134}Cs/^{137}Cs activity ratio of 0.56 in the trap samples reveals that their activity is unequivocally due to Chernobyl fallout and not to discharges of reprocessing plants of Sellafield, UK, or La Hague, France, which discharge Cs with activity ratios of 0.08 (in 1984) and 0.17 (in 1982), respectively[10].

The first activity occurred in cup N2 (open 14:53 4 May to 05:46 16 May). Because the centre of the Chernobyl cloud crossed the North Sea on 8 May[1] and no activity could have arrived before 3 May, massive transfer of radioactive isotopes to sediments started within ten days after fallout. Because of the high total flux in cup N2, it also yielded the highest total activity (50 Bq) of all the cups. The highest total specific activity (670 Bq g^{-1}) was found in cup N3 (17 Bq total), which received an amount of sediment about a factor of ten lower than cup N2 with the subsiding plankton bloom. Specific activities of ^{137}Cs, ^{134}Cs, ^{106}Ru and ^{103}Ru reached their highest values in cup N3, and ^{144}Ce, ^{95}Nb and ^{95}Zr reached their highest specific activities in cup N5, between 8 and 20 June. Total activities decreased from cup N3 to N9 but increased again from cup N9 to N13. Because these cups caught only very small amounts of sediment, the results have large error bars. Cone-sample N14 was checked also for activity on 16 April 1987. It contained 0.52 Bq g^{-1} (±10%) ^{137}Cs, ≤0.21 Bq g^{-1} ^{134}Cs, 45 (±35%) Bq g^{-1} ^{103}Ru, 2.4 (±27%) Bq g^{-1} ^{106}Ru, 2.8 (±13%) Bq g^{-1} ^{144}Ce and 11.6 (±22%) Bq g^{-1} ^{95}Nb (calculated from transient equilibrium with ^{95}Zr) (7,600 min counting time, activities standardized to 1 May 1986). These relatively low concentrations (compare with Table 2) indicate that either the cone material settled at the beginning or at the very end of the deployment period.

It is remarkable that in October short-lived nuclides such as ^{103}Ru and ^{141}Ce were still detectable even though 4.4 and 5.4 half-lives had passed. On 5–6 May 1986, a ^{103}Ru/^{137}Cs activity ratio of 2.0 was measured on air filters from Hamburg-Sülldorf[5]. The respective ratios of the trap samples range from 37 to 67 suggesting an adsorption of ^{103}Ru to marine particulate matter that is 20–30 times more effective than that of ^{137}Cs. Similarly, in all samples the activity of ^{106}Ru exceeded that of ^{137}Cs, even though the activity of ^{137}Cs was about a factor of 2–4 higher in the first fallout than the activity of ^{106}Ru. Partition coefficients for fractionation of Cs and Ru between sediments and water, however, suggest that Cs has a higher adsorption tendency than Ru. With respect to marine organisms the concentration factors are reversed[11]: that is, Ru concentrates in phytoplankton 200,000 times; in zooplankton, 30,000; in molluscs, 2,000; in crustaceans 100 and in fish, 10 times. The same factors for Cs are 20, 30, 30, 30 and 100 times, respectively. This leads to the conclusion that Ru was adsorbed mainly to the organic fraction of the settled sediment. Because the flux of organic matter is highly correlated with total flux ($r = 0.92$) but terrigenous flux is less significantly correlated with total flux ($r = 0.73$) and specific activities always show higher correlations with organic matter flux than with terrigenous flux (never statistically significant because of small number of samples), this inference is substantiated. In Fig. 2*b* the specific activities with respect to organic carbon (C_{org}) normalized to cup N3 = 100% are plotted. The pattern of concentration changes is very similar for Cs and Ru isotopes. Curves correlate significantly with each other ($r >$ 0.97), suggesting similar adsorption changes of Cs and Ru to organic substance with time. However, ^{144}Ce displays a different curve having a maximum in N5 where the other curves show minima.

The increasing activity from cup N9 to N13 may be explained by the arrival of a more contaminated water body at the trap

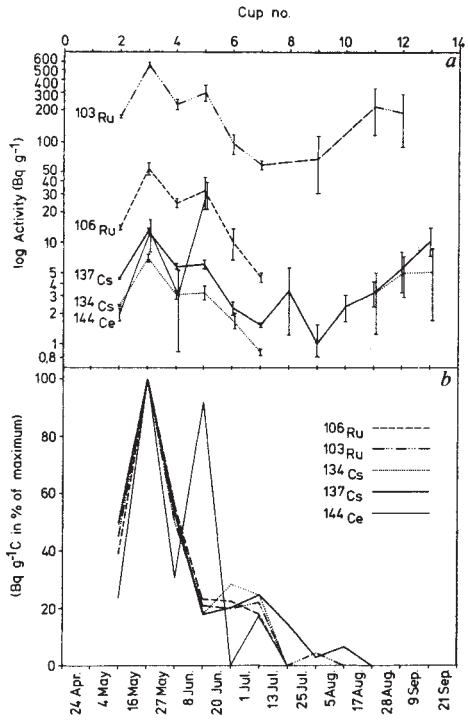


Fig. 2 *a*, Specific activity of sediment trap material with regard to five radioactive isotopes. *b*, Specific activity of sediment trap material normalized to organic carbon and to cup N3 = 100%.

site. Such patchiness has been demonstrated for the North Sea[5] and was caused by local rainfalls. The highest fallout activity measured was ~14 kBq m^{-2} ^{137}Cs at 57°20′ N, 03°00′ E. In two to three months this water mass could have moved to the trap site where original ^{137}Cs deposition may have been ~4 kBq m^{-2}. Of this, 30 Bq m^{-2} ^{137}Cs or 0.5-1% were transferred by particles to the sea floor during the observation period together with 17 Bq m^{-2} ^{134}Cs, 1,120 Bq m^{-2} ^{103}Ru, 99 Bq m^{-2} ^{106}Ru and 19 Bq m^{-2} ^{144}Ce.

In a May-June 1986 survey, between 0.07-0.3 Bq ^{137}Cs were measured per kg of seawater in the North Sea mixed layer[5]. Because $^{103}Ru/^{137}Cs$ and $^{131}I/^{137}Cs$ ratios of aerosols from Hamburg were two and six, respectively[5], total initial activity of seawater can be estimated us 0.7-3 Bq kg^{-1} due to ^{131}I, ^{137}Cs, ^{134}Cs, ^{103}Ru, and ^{106}Ru. The regulation laid down by the Commission of the European Community (CEC) allows a maximum activity of 600 Bq kg^{-1} for human food. For comparison, the total specific activity of particulates amounted to 670,000 Bq kg^{-1} in cup N3 (recalculated for 1 May and excluding ^{131}I which was not measurable in trap sediments), 1,100 times greater than the CEC maximum allowable contamination and up to 1,000,000 times higher than in contaminated seawater. Thus the marine food chain was burdened with high radioactive contamination for several weeks in May and June 1986. Interestingly, crustacean exuvia in the samples in the size fraction >1 mm did not show any measurable activity. Adsorption to particles and subsequent flocculation and sinking of these particles from the photic layer apparently scavenges the surface waters of nuclides but transfers the radioactive burden to the benthos. In the North Sea, the bulk flux of Chernobyl nuclides occurred during the spring plankton bloom forming a monoseasonal mark in the sedimentary record.

This work was funded by the German Minister for Research and Technology. Responsibility for the contents of this paper is with the authors. We acknowledge the assistance of I. Goroncy (DHI) in measuring gamma-activity spectra.

Received 17 April; accepted 4 September 1987.

1. Smith, F. B. & Clark, M. J. *Nature* **322**, 690-691 (1986).
2. Hohenemser, C. *et al. Nature* **321**, 817 (1986).
3. Thomas, A. J. & Martin, J. M. *Nature* **321**, 817-819 (1986).
4. Pringle, D. M., Vermeer, W. J. & Allen, K. W. *Nature* **321**, 569 (1986).
5. Deutsches Hydrographisches Institut. Die Auswirkungen des Kernkraftwerkes von Tschernobyl auf Nord- und Ostsee. *Meeresk. Beob. Ergebn. Dt. Hydrogr. Inst.* **62**, (1987).
6. Geyh, M. A., Hennig, G., Kruck, W. *et al. Bundesforschungsanstalt f. Landeskunde und Raumordnung 1987* (1/2) (1987).
7. Honjo, S., Connel, J. F. & Sachs, P. L. *Deep-Sea Res.* **27**, 745-754 (1980).
8. Gesellschaft für Reaktorsicherheit. *Neuere Erkenntnisse zum Unfall im Kernkraftwerk Tschernobyl, GRS-S-40*, 2nd edn., 74 pp. (1987).
9. Backhaus, J. O. *Estimates of the Variability of Low Frequency Currents and Flushing-times of the North Sea.* Vol. C24 (International Council for the Exploration of the Sea, 1984).
10. Kautsky, H. *Deut. Hydrogr. Zeitschr.* **40**, 49-69 (1987).
11. *Review of the Continued Suitability of the Dumping Site for Radioactive Waste in the North-East Atlantic* (Nuclear Energy Agency, Paris, 1985).

Male predation risk determines female preference in the Trinidad guppy

Felix Breden & Gregory Stoner

Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211, USA

The 'runaway process of sexual selection' explains the evolution of extreme sexual behaviours and adornments and requires that a preferred secondary sexual character expressed in one sex be genetically correlated with a preference for this character expressed in the other sex[1-4]. Because of this genetic correlation, individuals selecting the preferred character also select higher levels of preference, producing rapid, self-reinforcing evolution of both characters. Preference for increased development of the secondary sexual character is counterbalanced at equilibrium by natural selection acting against this character. Variation in predation against brightly coloured male guppies, *Poecilia reticulata*, presents a unique opportunity to test a prediction of this counterbalance between sexual and natural selection, namely that geographical variation in selection against the preferred character should lead to parallel variation in preference[5]. Here we present evidence that female guppies from high-predation populations show a genetically determined, lower preference for brightly coloured models of male guppies than do females from areas of low predation.

Laboratory and field experiments have shown that the decreased brightness of adult male guppies in high-predation populations is caused by direct selection against bright males by other fish species[6-9]. A previous study demonstrated phenotypic differences between females from high-predation areas, exhibiting reduced preference for bright males, and females from low-predation areas. In this letter we show that these phenotypic differences are caused by genetic differentiation between these populations.

Guppies were collected in August 1986 from two high-predation and two low-predation populations in the Northern Range of mountains in Trinidad, West Indies (Table 1). There is one small cyprinidont that preys on guppies in the low-predation populations, while the high-predation populations experience several species of predatory fish Several authors have described these populations in detail[10-12]. The high- and low-predation populations are separated by less than 5 km. Stocks were established from approximately 20 adult females and 10 adult males per population, maintained at 25 °C on a 14:10 L:D photoperiod, and fed fish flakes *ad libitum*. Ten wild-caught females were removed from the low-predation and high-predation stocks (these could have been inseminated by males in the field and in the stock populations) and isolated in 40-l aquaria. Within 24 h of parturition, 6-8 offspring were removed from each brood and transferred to a 20-l aquarium with the same environmental

Table 1 Average choice scores for sibships of females from high- and low-predation populations

High-predation treatment						Low-predation treatment					
River	Population	Female	Sibship 1	Sibship 2	Average	River	Population	Female	Sibship 1	Sibship 2	Average
Aripo	Lower	APLN1	−0.02	0.41	0.20	Arima	Tripp	AMTN3	0.35	-	0.35
Aripo	Lower	APLN2	−0.35	-	−0.35	Arima	Tripp	AMTN4	0.20	-	0.20
Aripo	Lower	APLN3	−0.03	-	−0.03	Arima	Tripp	AMTN6	0.35	−0.03	0.16
Aripo	Lower	APLN4	−0.05	-	−0.05	Arima	Tripp	AMTN9	0.12	-	0.12
Aripo	Lower	APLN5	−0.59	−0.17	−0.38	Arima	4M	AM4MN2	0.25	-	0.25
Aripo	Lower	APLN7	−0.48	-	−0.48	Arima	4M	AM4MN3	0.27	0.48	0.38
Arima	Lower	AMLN3	−0.12	-	−0.12	Arima	4M	AM4MN5	0.17	0.58	0.38
Arima	Lower	AMLN5	−0.36	0.01	−0.18	Arima	4M	AM4MN6	0.53	-	0.53
Arima	Lower	AMLN6	−0.22	-	−0.22	Arima	4M	AM4MN9	0.18	-	0.18
Arima	Lower	AMLN7	−0.20	-	−0.20	Arima	4M	AM4MN10	0.25	-	0.25

Naive females were tested for preference for bright or drab male guppy models in a 40-l aquarium, divided into three compartments. The length of the two end compartments was 12.5 cm and the middle compartment was 25 cm; the water level was 25 cm. The outside glass of the middle compartment was divided into a 5×5 grid of 5-cm squares. An observer sat 1 m from the grid. Male guppy models were carved from balsa wood and suspended 7 cm above the substrate by monofilament wire anchored to the bottom of the tank. Each model was 4.8 cm total length by 1.3 cm maximum height (approximately twice as long as males used in our previous study[10] and painted dull yellow. To mimic bright males collected from low-predation stocks and used in our previous study[10], we painted 56% of the bright model with a combination of black, bright yellow, orange and blue spots. Aerators were placed in the compartments with the models, with the aerator nearer the bright model, producing a higher degree of motion. For each sibship, the positions of the males were first determined at random, and then alternated between trials. Female choice was determined by noting the position of the female relative to the drab and bright models every 30 s for a 10-min interval. Observations showed that when a female was in the column of squares nearest an end compartment she was almost always oriented toward that compartment; this suggested the following scoring procedure that gave greatest weight to being in the column of squares nearest the male[10]. If a female was in the column of squares nearest the bright model at the beginning of an interval, her preference score for that model (N_B) increased two points; if she also remained in that column for the entire interval, N_B increased a total of three points. If a female was within the column of squares between the centre column and the column adjacent to the bright model, N_B increased one point. A female's choice score was calculated as $(N_B - N_D)/(N_B + N_D)$, where N_D is the females's score relative to the drab model. Positive values indicate choice of the bright model. Values of $(N_B + N_D)$, representing total activity per female, ranged from 23 to 50, and showed no significant variation at any level of the nested analysis of variance. Sibship 1 and 2 refer to broods removed sequentially from one female.

conditions as above except that diet was supplemented by brine shrimp nauplii. Males were removed when they were 28 days old, before the development of pigmentation. At 15 mm standard length (approximately day 40), three females per sibship were tested for preference for a bright and drab model. Models were used to remove variation due to differences between the activity of males within the same level of brightness or differences between the activity of the same males used in several trials. Trials were conducted in a 40-l aquarium divided into three compartments by glass. Each female in a sibship was placed sequentially in the middle compartment and presented with a bright and dull male guppy model, one in each of the end compartments. After a 15-min equilibration period, the preference of each female was determined by recording her position relative to the bright or drab model at 30-s intervals for 10 min (see Table 1 legend for a description of the models and the method of calculating choice scores).

Table 1 presents the average choice scores per sibship, and the per female averages for those females that had two sibships tested. All mean scores for low-predation females indicate preference for the bright model, and all but one high-predation female average showed preference for the drab model. A nested analysis of variance[13] showed a significant effect of predation regime on preference (Table 2, $P < 0.05$). Because offspring were raised in identical conditions, differences among females reflect genetic differences[14]. Maternal effects on behaviour were minimized by removing offspring from the maternal tank within 24 h. The assumption that the behaviour elicited by the models represents behaviour similar to female choice of real males is supported by the fact that these females performed similarly to those tested with live males in our study of phenotypic differentiation[10], including directed movements toward the models, which is considered a female sexual response in guppies[15]. The discrimination of high-predation females against bright males could be caused by a direct decrease in survivorship experienced by females in high-predation areas that prefer to mate with bright males, perhaps associated with being in the vicinity of bright males at a high risk of predation. However, this risk is likely to be small, because females are sexually receptive for only a few days after parturition and do not spend much time actively courting.

Although there has been much theoretical work on the evolution of female choice and male sexual characters by the process of sexual selection, few theoretical studies have investigated the importance of geographical variation in natural selection on the evolution of female choice and little empirical evidence has been gathered to validate these models. The most important question in determining what model of sexual selection is appropriate for the system we studied is whether geographical differences in the level of predation cause differences in the optimum viability phenotype for males, in the intensity of selection around this optimum, or in both[10]. If all levels of brightness are detrimental in this system, then the optimum viability phenotype in both high- and low-predation areas would be a zero level of brightness. This would require modification of several models of sexual selection that assume a distribution of phenotypes and a selection function symmetrical around the optimum viability phenotype[2,4], an impossibility if this optimum were zero.

Our main result is that these populations of guppies are genetically differentiated for female preference, with most females from high-predation populations preferring the drab model and most low-predation females preferring the bright model. The fact that only two of the high-predation sibships showed preference for the bright model and only one low-predation sibship showed preference for the drab model suggests that there is little genetic variation within these populations for female preference. Differentiation has occurred among these populations even though direct selection acting on female preference is probably minimal. Thus we have shown that female preference has evolved as a correlated response to direct selection acting against bright male offspring produced by females mating with bright males in high-predation areas. Although several studies have shown female preference for increased

Table 2 Nested analysis of variance for female choice scores

Source	Degrees of freedom	Sum of squares	*F* ratio	(*P*)
Predation treatment	1	3.625	48.97	0.020
Population	2	0.148	0.62	0.551
Female	16	1.917	0.56	0.832
Sibship	6	1.276	1.88	0.102
Error	52	5.882	—	—

development of certain male characters[15-18], and that within-population variation in female preference has a genetic basis[19,20], this is the first study showing between-population genetic variation in female preference, responding as a character genetically correlated with a male secondary sexual character. The amount of genetic differentiation is similar to that observed in a study of schooling behaviour in the Trinidad guppy[12], confirming that behaviour patterns of vertebrate populations can respond evolutionarily to fine-scale geographical variation in selection.

We thank Janice Winters and Carol Kelly for field assistance, the University of Missouri Graduate Research Council for support, the Trinidad Ministry of Agriculture, Lands and Fisheries for permission to export fish, and Mark Davis, Carl Gerhardt, Glenn Hausfater and Jamie Scott for comments.

Received 3 July; accepted 7 September 1987.

1. O'Donald, P. *Genetic Models of Sexual Selection* (Cambridge University Press, 1980).
2. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3721-3725 (1981).
3. Kirkpatrick, M. *Evolution* **36**, 1-12 (1982).
4. Heisler, I. L. *Evolution* **38**, 1283-1295 (1984).
5. Lande, R. *Evolution* **36**, 213-223 (1982).
6. Endler, J. A. *Evol. Biol.* **11**, 319-364 (1978).
7. Endler, J. A. *Evolution* **34**, 76-91 (1980).
8. Endler, J. A. *Evolution* **36**, 178-188 (1982).
9. Endler, J. A. *Envir. Biol. Fishes* **9**, 173-190 (1983).
10. Stoner, G. & Breden, F. *Behav. Ecol. Sociobiol.* (in the press).
11. Reznick, D. & Endler, J. A. *Evolution* **36**, 160-177 (1982).
12. Breden, F., Scott, M. & Michel, E. *Anim. Behav.* **35**, 618-620 (1987).
13. *SAS User's Guide: Statistics* (SAS Institute, Cary, North Carolina, 1982).
14. Reznick, D. *Evolution* **36**, 1236-1250 (1982).
15. Houde, A. *Evolution* **41**, 1-10 (1987).
16. Andersson, M. *Biol. J. Linn. Soc.* **17**, 375-393 (1982).
17. Sullivan, B. K. *Anim. Behav.* **31**, 1011-1017 (1983).
18. Bischoff, R. J., Gould, J. L. & Rubenstein, D. I. *Behav. Ecol. Sociobiol.* **17**, 253-255 (1985).
19. Majerus, M. E. N., O'Donald, P. & Weir, J. *Nature* **300**, 521-523 (1982).
20. Heisler, I. L. *Genet. Res.* **44**, 133-149 (1984).

Cytoplasmic calcium regulates voltage-dependent ion channels in plant vacuoles

R. Hedrich & E. Neher

Department of Membrane Biophysics, Max-Planck-Institut für biophysikalische Chemie, D-3400 Göttingen, FRG

Vacuoles, the storage organelles of plants, can constitute up to 90% of the cell volume. Transport of solutes such as sugars and organic and inorganic ions across the vacuolar membrane constitutes one of the major cellular achievements during diurnal and annual cycles of carbon metabolism. Recent investigations by patch-clamp techniques[1] identified ion channels and pumps as pathways for the movement of ions and metabolites[2-4]. The mechanisms by which these transport proteins are regulated remain unknown. Changes in the cytoplasmic concentration of Ca^{2+} have been postulated to be involved in the physiology of various ion transport processes in plant cells[5-9]. But in contrast to animal cells[10], to date no direct evidence exists for the modulation of ion-transport proteins by Ca^{2+} in plant cells. Here we report that in plant vacuoles, changes in cytoplasmic Ca^{2+} regulate the activation of two distinct types of voltage-dependent ion channels. A

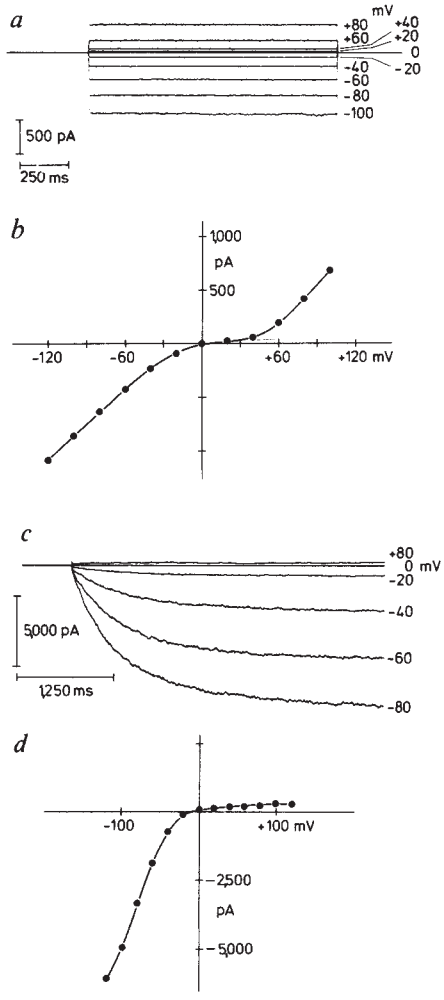


Fig. 1 *a*, Whole-vacuole currents in response to a series of voltage steps from a holding potential of 0 mV with the vacuole exposed to 10^{-8} M Ca^{2+} on the cytoplasmic side of the membrane. Solutions contained: 100 mM KCl, 5 mM Tris-MES *p*H 7.2, 2 mM $MgCl_2$, buffered at 1 mM total EGTA in the bath; 10 mM KCl, 5 mM MES/Tris *p*H 5.5, 2 mM $MgCl_2$, 1 mM $CaCl_2$ in the pipette. *b*, Current-voltage relation (*I-V*) of vacuolar currents as shown in *a*. *c*, Membrane currents during a voltage-pulse sequence recorded with 10^{-4} M $CaCl_2$ in the bath. The membrane potential was stepped from a holding potential of 0 mV to various potentials ranging from −80 to +80 mV. Six current traces have been superimposed. During the voltage steps to hyperpolarized potentials (vacuole inside negative) slowly activating currents could be observed. Solutions contained: 100 mM KCl, 2 mM $MgCl_2$, 10^{-4} M $CaCl_2$, 5 mM Tris-MES *p*H 7.5 in the bath and 100 mM KCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$, 5 MES/Tris *p*H 5.5 in the pipette. *d*, *I-V* relation of the vacuolar currents as shown in *c*.

Methods. Sugarbeet vacuoles were isolated from fresh taproots grown in the field or greenhouse. A slice of storage tissue was cut off with a razor blade and the surface rinsed with some drops of bathing medium to wash the extruded vacuoles directly into the recording chamber[3]. The osmolarity of the root sap was measured with a vapour-pressure osmometer (Wescor 5100 C) so that solutions used for the experiments could be adjusted accordingly with D-sorbitol. Patch pipettes were sealed against a vacuole and the underlying membrane was broken by a pulse of +0.6 V and −0.6 V of 3 ms each. After access to the vacuole, the pipette solution equilibrated with the intravacuolar space (whole-vacuole). The ability of the electrogenic pump to produce ATP-induced inward currents[2,3] was used to document the viability of isolated vacuoles. Current measurements were made with an EPC-7 patch clamp (List Electronic), low-pass filtered at 1,000 Hz with an eight-pole Bessel characteristic. Data were stored on tape and analysed using a PDP-11/73 microcomputer.

60–80 pS channel appeared at Ca^{2+} levels above 10^{-7} M, activating slowly (SV—slow vacuolar type) upon application of hyperpolarizing voltage pulses. The voltage-threshold of appearance changed with the cytoplasmic Ca^{2+} concentration. At low Ca^{2+} a smaller channel (30–40 pS) activated instantaneously (FV—fast vacuolar type) upon application of both depolarizing and hyperpolarizing voltages. The two channel types modulated by physiological changes of cytoplasmic Ca^{2+} provide for versatile regulation of ion and metabolite fluxes between the cytoplasm and the vacuole.

The experiments were performed on mechanically isolated sugarbeet vacuoles[3]. Ionic currents were measured in the whole-vacuole configuration[2] (analogous to the 'whole-cell' configuration[1]) and from excised membrane patches. Figure 1*a* shows whole-vacuole currents in response to various potential steps from a holding potential of 0 mV at low (10^{-8} M) external Ca^{2+} (cytoplasmic side). With 10 mM KCl inside and 100 mM KCl outside the vacuole, hyperpolarizing voltages (vacuole inside negative) evoke inward currents (into the vacuole) and depolarizing potentials evoke outward currents (Fig. 1*a*). Because the currents elicited by the voltage steps show fast activation, they will be referred to as FV-type currents. The current–voltage relationship (I–V) in asymmetric KCl revealed a region of high resistance between 0 and +20 mV (Fig. 1*b*). The exact reversal potential was determined by the reversal of the single-channel currents.

Raising the extravacuolar Ca^{2+} concentration above 10^{-6} M resulted in a dramatic change of the current activation kinetics as well as the rectification properties. When the membrane potential was held at 0 mV, hyperpolarizing voltage pulses produced SV-type currents (Fig. 1*c*) as shown previously[2,3]. The I–V relationship in high (10^{-4} M) Ca^{2+} and symmetric K^+ solutions shows that SV-type currents were generated by hyperpolarizing potentials only, as indicated by a strong inward rectification (Fig. 1*d*). With high cytoplasmic Ca^{2+}, hyperpolarized potentials produced SV-type currents that were about 10 times larger than the FV-type currents seen with low Ca^{2+}.

Figure 2 illustrates concentration-dependent enhancement of SV-type currents by Ca^{2+}. A vacuole, containing only SV-type currents, was clamped to 0 mV and repetitively stepped to −80 mV while the cytoplasmic side was exposed to various Ca^{2+} concentrations. The increase in Ca^{2+} from 10^{-8} to 10^{-4} M led to the activation of SV-type currents (Fig. 2*a*). A plot of steady-state currents, derived from 4-s pulses to −80 mV, against extravacuolar Ca^{2+} concentration shows a sigmoidal dependence of the SV-type current on cytoplasmic Ca^{2+} (Fig. 2*b*). The activation curve of the SV-type current could be fitted by a Hill equation with an apparent K_d of 1.5×10^{-6} M Ca^{2+} and a Hill coefficient of 1, indicating a first-order reaction. As in animal cells, the gating of the Ca^{2+}-dependent channel depends both on Ca^{2+} concentration and on voltage[11]. When an analysis as in Fig. 2*b* was performed at several test potentials it was found that the threshold potential of activation shifted towards more depolarized values when cytoplasmic Ca^{2+} was increased (Fig. 2*c*).

In contrast to the SV-type current, the FV-type current is inactivated by high cytoplasmic Ca^{2+}. It was not observed at high cytoplasmic Ca^{2+} and was activated at concentrations lower than 10^{-6} M. The FV-type currents were not found in all vacuoles studied, but were consistently found in vacuoles from young sugarbeets. In some experiments the FV-type currents were labile, often showing a marked 'rundown' within 10–15 min. This prevented a detailed study of their Ca^{2+}-dependence and permeability sequence as shown for the SV-type currents (Fig. 2 and ref. 3).

To determine whether one or two different channel conductances were involved in FV- and SV-type currents outside-out patches[1] were used which enabled us to resolve the single channels which contribute to the macroscopic currents (Fig. 1*a* and *b*). Patches were held at −14 mV and stepped to various hyperpolarizing and depolarizing potentials during 1.5-s pulses.

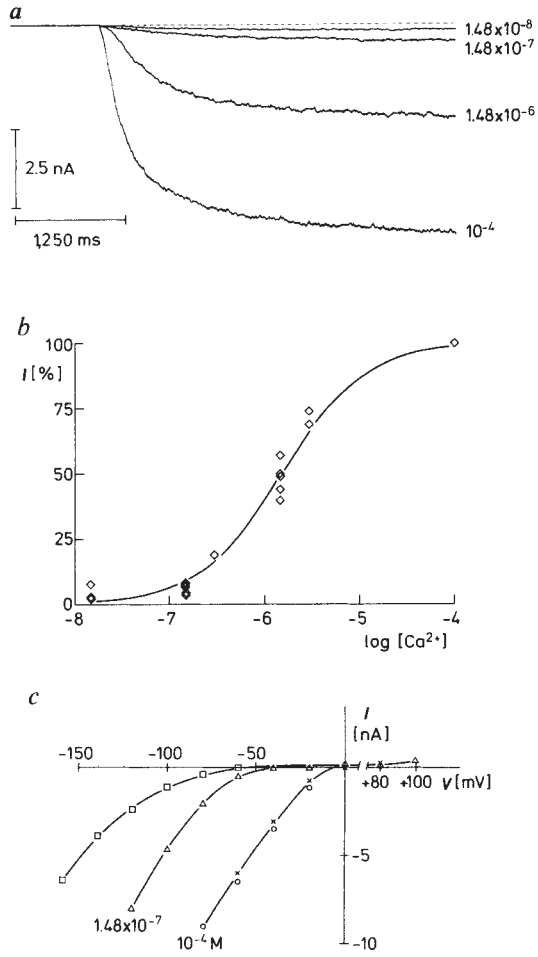


Fig. 2 *a*, Whole-vacuole currents in response to voltage pulses to −80 mV from a holding potential of 0 mV recorded in various Ca^{2+}/EGTA ratios. Four current traces recorded at free Ca^{2+} concentrations of 1.48×10^{-8} M, 1.48×10^{-7} M and 10^{-4} M in the extravacuolar medium were superimposed. Inward currents are activated by the rise in 'cytoplasmic' Ca^{2+}. Solutions contained: external, 200 mM KCl, 5 mM Tris-MES *p*H 7.2, 2 mM $MgCl_2$; internal (pipette), 200 mM KCl, 5 MES/Tris *p*H 5.5, 2 mM $MgCl_2$, 1 mM $CaCl_2$. All Ca^{2+}/EGTA buffers were computed assuming an apparent dissociation constant of EGTA of 1.48×10^{-7} (at *p*H 7.2)[28]. *b*, Current-Ca^{2+} relationship of the steady-state currents resulting from voltage pulses to −80 mV as shown in *a*. Results from five separate vacuoles are shown. Currents were normalized by setting steady-state currents measured in 10^{-4} M Ca^{2+} to 100%. The curve was fitted by the least-squares method using the Hill equation. The fitted curve had a Hill coefficient of 1 and a K_d of 1.5×10^{-6} M. *c*, Voltage dependence of the vacuolar currents in □, 1.48×10^{-8} M; △, 1.48×10^{-7} M and 10^{-4} M Ca^{2+}. The increase in the free Ca^{2+} concentration shifts the threshold potential for activation to more depolarized values.

In low cytoplasmic Ca^{2+} concentrations (10^{-8} M) hyperpolarizing potentials evoked inward channel currents and depolarizing voltages evoked outward channel currents (Fig. 3*a*). The unitary conductance of this channel is about 30–40 pS in symmetrical 200 mM KCl solutions. The voltage dependence as well as the fast activation are in good agreement with the whole-vacuolar FV-type currents (Fig. 1*a*) found under identical conditions.

Increasing Ca^{2+} concentration to 10^{-4} M causes the FV-type channel to disappear, and a 60–80 pS channel to become prominent. Channels with similar properties could be observed in 'vacuole-attached' (analogous to 'cell-attached'[1]) patches with pipette solutions containing 1 mM $CaCl_2$. The 60–80 pS channels are elicited only by hyperpolarizing voltages and have slow activation kinetics. The time course of the single-channel events

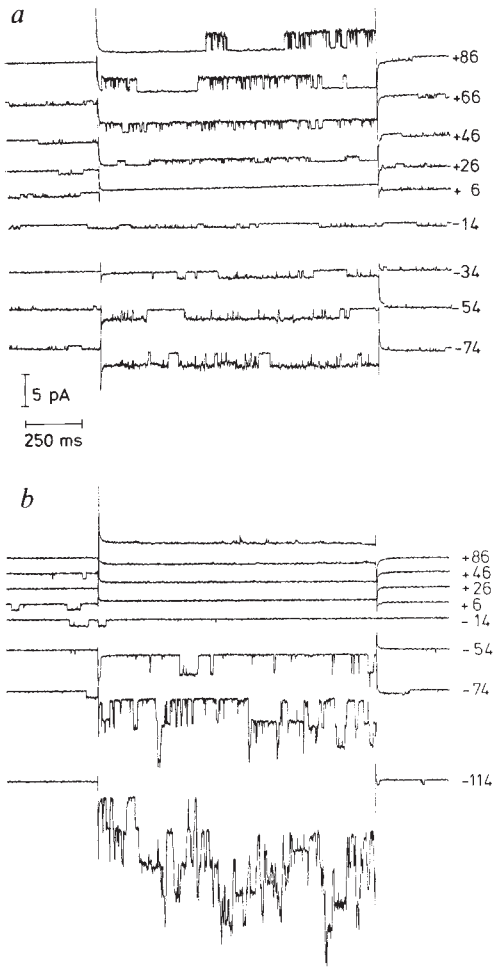


Fig. 3 *a*, Currents recorded from an outside-out patch incubated in 10^{-8} M Ca^{2+} on the cytoplasmic side of the membrane. The patch was held at −14 mV and during a 1.5-s pulse stepped to various positive and negative potentials as indicated. Hyperpolarizing potentials evoked inward single-channel currents (into the vacuole) and depolarizing voltages evoked outward currents which indicate a conductance of 30 pS. These channels were activated concomitantly with the onset of the voltage pulse. The voltage dependence and kinetic behaviour of this channel are consistent with the macroscopic FV-type currents measured in low Ca^{2+} shown in Fig. 1*a*. *b*, Single-channel openings recorded from the same patch after perfusion with 10^{-4} M $CaCl_2$ using the same voltage protocol as in *a*. Hyperpolarizing potentials from a holding potential of −14 mV evoked channels with a conductance of 60 pS. The channels were mainly activated during the 1.5-s voltage pulse. However, a few channel openings were also observed at the baseline because a holding potential of −14 mV partially activates the channel. Depolarizing voltages elicited no channel openings. The slow activation and voltage dependence of the 60 pS channel reflect the macroscopic SV-type currents in high Ca^{2+} shown in Fig. 1*b*. Scale as in *a*. Solutions in *a* and *b* contained: 200 mM KCl, 5 mM Tris-MES *p*H 7.2, 2 mM $MgCl_2$ in the bath and 200 KCl, 5 mM MES-Tris *p*H 55, 2 mM $MgCl_2$, 1 mM $CaCl_2$ in the pipette.

shown in Fig. 3*b* was consistent with the slow activation of the macroscopic currents (Fig. 1*b*), indicating that at high Ca^{2+}, vacuolar currents are carried by SV-type channels. It should be emphasized that at high Ca^{2+} FV-type channels were never observed ($N > 20$), despite the fact that they should be very prominent in the positive voltage range, where SV-type channels are not activated.

In excised patches the permeability ratio P_{K^+}/P_{Cl^-} for both channel types was calculated to be ~6, as derived from reversal potentials between −25 and −30 mV of single-channel currents with 20 mM KCl in the extravacuolar solution and 200 mM KCl in the vacuole. For the SV-type channel a selectivity sequence $P_{K^+} = P_{Na^+} > P_{Ac^-} > P_{NO_3^-} > P_{Mal^{2-}} > P_{Cl^-}$ was found[3], where Ac^- indicates acetate ion, and Mal^{2-} malate ion.

That vacuoles containing only one type of channel could be found under certain conditions points to the existence of two separate channels or else to two profoundly different forms of the same channel, rather than a reversible interconversion between FV and SV-type channels brought about by changes in Ca^{2+} concentration.

These data present the first demonstration of Ca^{2+} modulated channels in higher plant cells. Two channel types with inverse sensitivity to Ca^{2+} and different voltage-dependence were found in isolated sugarbeet vacuoles. Changes in the physiological range of the cytoplasmic Ca^{2+} concentration (10^{-7}–10^{-5} M)[12-14] lead to pronounced changes in the conductance properties of the vacuolar membrane due to the activation or suppression of SV- or FV-type channels. In previous studies[2], which were performed at high Ca^{2+} concentrations and thus were concerned with SV-type channels, it was found that the permeability sequence for ions of physiological relevance corresponds to that found by tracer flux measurements on vacuole suspensions. This can be taken as evidence that the SV-type channel is a pathway for the movement of organic and inorganic ions across the vacuolar membrane[2,3].

The activity of a proton pump, regulated by the adenylate ratio[15], and thus reflecting the energy status of the cell[16] was previously shown to depolarize the vacuolar membrane[2,3]. Under these conditions and with cytoplasmic Ca^{2+} being low, only FV-type channels should be active. They can provide a pathway for anions which accumulate at positive vacuole potential and allow the equilibration of K^+ (ref. 17). Elevation of cytoplasmic Ca^{2+}, following diverse physiological stimuli such as light-dark transitions[12], (H. Felle, personal communication), inhibition of calmodulin[18] or phosphoinositol metabolites[19-21], and reduction of the pump-generated positive potential would open the SV-type channel for the release of anions assuming that the Ca^{2+}-dependence shown here prevails under these conditions. Experiments on stomatal movement provide evidence that high Ca^{2+} prevents the accumulation of anions and enhances their release[22-24]. The release of anions and accompanying counterion fluxes are used by plant cells to adjust the osmotic pressure. In addition organic anions (malate) released from the vacuole, serve as substrates for cytoplasmic carbon metabolism[25-27]. The properties of the two types of channels described are thus consistent with their likely functions in solute transport.

The Ca^{2+}-dependence of vacuolar channels as reported here in storage tissue was also found in vacuoles from photosynthetic and phototropic tissue (data not shown) and may therefore provide a general mechanism for the regulation of ion transport in plant vacuoles.

We thank J. I. Schroeder, R. Penner, W. Lahr, F. Ashcroft and M. B. Jackson for critical reading of the manuscript.

Received 29 July; accepted 7 September 1987.

1. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
2. Hedrich, R., Flügge, U. I. & Fernandez, J. M. *FEBS Lett.* **3910**, 228-232 (1986).
3. Coyaud, L., Kurkdjian, A., Kado, R. & Hedrich, R. *Biochim. biophys. Acta.* (in the press).
4. Kolb, H-A., Köhler, K. & Martinoia, E. J. *Membrane Biol.* **95**, 163-169 (1987).
5. Hepler, P. K. & Wayne, R. O. *A. Rev. Pl. Physiol.* **36**, 397-439 (1985).
6. Marme, D. & Dieter, P. in *Calcium and Cell Function* Vol. IV (ed. Cheung, W. Y.) 263-311 (Academic, New York, 1983).
7. Jones, R. L., Deikman, J. & Melroy, M. J. in *Molecular and Cellular Aspects of Calcium in Plant Development* (ed. Trewavas, A. J.) 175-183 (Plenum, New York, 1986).
8. Lunevsky, V. Z., Zherelova, O. M., Vostrikov, I. Y. & Berestovsky J. *Membrane Biol.* **72**, 43-58 (1983).
9. Mimura, T. & Tazawa, M. *Protoplasma* **118**, 49-55 (1983).
10. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, Massachusetts, 1984).
11. Marty, A. *Trends Neurosci.* **6**, 262-265 (1983).
12. Miller, A. J. & Sanders, D. *Nature* **326**, 397-400 (1987).
13. Williamson, R. E. & Ashley, C. C. *Nature* **296**, 647-651 (1982).
14. Brownless, C. & Wood, J. W. *Nature* **320**, 624-626 (1986).
15. Bennett, A. B. & Spanswick, R. M. *Pl. Physiol.* **74**, 545-548 (1984).

16. Gardeström, P. *FEBS Lett.* **4412**, 114-118 (1987).
17. Martinoia, E., Flügge, U. I., Kaiser, G., Heber, U. & Heldt, H. W. *Biochim. biophys. Acta* **806**, 311-319 (1985).
18. Gilroy, S., Hughes, W. A. & Trewavas, A. J. *FEBS Lett.* **4436**, 133-137 (1987).
19. Rincon, M. & Boss, W. F. *Pl. Physiol.* **83**, 395-398 (1987).
20. Schumaker, K. S. & Sze, H. *J. biol. Chem.* **262**, 3944-3946 (1987).
21. Morse, M. J., Crain, R. C. & Satter, R. L. *Pl. Physiol.* **80**, 486 (1986).
22. DeSilva, D. L. R., Hetherington, A. M., & Mansfield, T. A. *New Phytol.* **100**, 473-483 (1985).
23. DeSilva, D. L. R., Cox, R. C., Hetherington, A. M. & Mansfield, T. A. *New Phytol.* **101**, 555-563 (1985).
24. Schwartz, A. *Pl. Physiol.* **79**, 1003-1005 (1985).
25. Schnabl, H., Elbert, C. & Krämer, G. *J. exp. Bot.* **33**, 996-1003 (1983).
26. Gerhardt, R. & Heldt, H. W. *Pl. Physiol.* **75**, 542-547 (1984).
27. Lüttge, U. *New Phytol.* **105**, 593-629 (1987).
28. Martell, A. E. & Smith, R M. *Critical Stability Constants* (Plenum, London, 1974).

cDNA cloning of bovine substance-K receptor through oocyte expression system

Yasuo Masu, Kazuhisa Nakayama, Hiromi Tamaki, Yoshio Harada*, Motoy Kuno* & Shigetada Nakanishi

Institute for Immunology and * Department of Physiology, Kyoto University Faculty of Medicine, Kyoto 606, Japan

The neuropeptide receptors which are present in very small quantities in the cell and are embedded tightly in the plasma membrane have not been well characterized. Mammals contain three distinct tachykinin neuropeptides, substance P, substance K and neuromedin K, and it has been suggested that there are multiple tachykinin receptors[1,2]. By electrophysiological measurement, we have previously shown that *Xenopus* oocytes injected with brain and stomach mRNAs faithfully express mammalian substance-P and substance-K receptors, respectively[3]. Here we report the isolation of the cDNA clone for bovine substance-K receptor (SKR) by extending this method to develop a new cloning strategy. We constructed a stomach cDNA library with a cloning vector that allowed *in vitro* synthesis of mRNAs and then identified a particular cDNA clone by testing for receptor expression following injection of the mRNAs synthesized *in vitro* into the oocyte system. Because oocytes injected with exogenous mRNAs can express numerous receptors and channels, our new strategy will be applicable in the general molecular cloning of these proteins. The result provides the first indication that the neuropeptide receptor has sequence similarity with rhodopsin-type receptors (the G-protein-coupled receptor family) and thus possesses multiple membrane-spanning domains.

The principle of the new cloning strategy is illustrated in Fig. 1. A mixture of complementary DNAs was first synthesized from bovine stomach messenger RNA by using a specific poly(dT) primer. This primer contains multiple restriction sites next to the poly(dT) tail so that when an appropriate enzyme is chosen, cleavage of one of these sites terminates the subsequent *in vitro* transcription of cDNAs cloned in λ immediately after the poly(A) tail but leaves the receptor mRNA sequence intact. After synthesis of single- and double-stranded cDNAs, the resultant cDNA mixture was integrated into the *Eco*RI site located downstream from the SP6 promoter in the λ vector DNA. We chose λ as a cloning vector, because it is capable of incorporating a large DNA insert, increasing the proportion of clonal cDNAs encoding whole mRNA sequences. Appropriate numbers of λcDNA clones were plated after *in vitro* packaging[4] and the clonal λcDNA mixture was extracted and cleaved with an appropriate restriction enzyme. The resultant λcDNA mixture was transcribed *in vitro* by specific SP6 RNA polymerase in the presence of the capping nucleotide[5]. The mRNA mixture synthesized was injected into oocytes which were then tested for the expression of the receptor by measuring electrophysiological response to application of substance K. Following our experimental design, a mixture of cDNA clones comprising the receptor cDNA sequence would be identified by detecting the electrophysiological response of oocytes injected with the mRNA to application of substance K. The receptor cDNA clone could subsequently be purified by stepwise fractionations of the response-evoking cDNA mixture on the basis of the *in vitro* mRNA synthesis and electrophysiological measurement.

We began by examining SKR expression for the λcDNA mixture comprising $\sim 3\times10^5$ cDNA clones. A clear response to substance K was observed for oocytes injected with the mRNA derived from the cDNA mixture. We therefore attempted to purify the cDNA clone for SKR by following the ability to evoke a response through stepwise fractionations of the cDNA clone mixture. Testing fractions containing ~50,000, ~20,000, ~2,000 and 83 λcDNA clones, we finally obtained a single clone (λSKR56) that was capable of inducing electrophysiological response to substance K. When an *Eco*RI fragment containing the whole cDNA sequence of λSKR56 was inserted immediately downstream from either the SP6 promoter or the T7 promoter in the plasmid DNA (pSKR56S or pSKR56T), the mRNA synthesized *in vitro* by SP6 RNA polymerase (pSKR56S) or T7 RNA polymerase (pSKR56T) again induced an electrophysiological response to substance K in the oocyte expression system (see Fig. 4). Thus, we concluded that this cDNA sequence is sufficient for the induction of an electrophysiological response to substance K. The SKR cDNA was sequenced by the chain termination method[6].

Figure 2 shows the 2,458-nucleotide sequence of the cloned cDNA and the amino-acid sequence deduced for SKR, which was assigned from the longest open-reading frame of the cDNA sequence. This polypeptide consists of 384 amino-acid residues (relative molecular weight (M_r) = 43,066). The hydropathicity profile[7] of the deduced amino-acid sequence and a comparison of its sequence with those of other proteins[8] revealed seven hydrophobic segments and significant sequence similarity shared with rhodopsin-type receptors[9-15]. Figure 3 shows the alignment of the amino-acid sequences for SKR and rhodopsin-type receptors (β_2-adrenergic receptor, βAR; muscarinic acetylcholine receptors, M_1AChR and M_2AChR; and opsin). The degree of sequence identity is 24, 24, 22 and 21% for SKR-βAR, SKR-M_1AChR, SKR-M_2AChR and SKR-opsin comparisons, respectively.

Rhodopsin-type receptors have a structure consisting of seven hydrophobic membrane-spanning domains with an extracellular amino-terminus and a cytoplasmic carboxy-terminus[9,16,17]. A similar transmembrane model for SKR, with seven membrane-spanning segments, can be suggested from its primary structure (Fig. 3). The seven putative transmembrane α-helices of SKR consist of a continuous stretch of 20-24 uncharged amino-acid residues, except that segments II, V and VI contain Asp 79, His 198 and His 267, respectively. The residue corresponding to Asp 79 is conserved for all rhodopsin-type receptors. Several cysteine residues (corresponding to residues 106, 181, 262, 309 and 324) are also conserved in all or most of the receptors. In contrast to the putative transmembrane domains, the amino- and carboxy-terminal regions and the third cytoplasmic loop (loop V-VI) differ in length and in the amino-acid sequence between SKR and the rhodopsin-type receptors. However, these regions of SKR also show a pattern similar to the rhodopsin-type receptors[17], including potential *N*-glycosylation sites[18] (Asn 11 and Asn 19) in the amino-terminal region with no signal sequence[19], and the presence of many Ser and Thr residues in the carboxy-terminal region as possible phosphorylation sites[17]. Thus, the close similarities between SKR and rhodopsin-type receptors, in both their amino-acid sequences also general structural features, indicate that SKR belongs to the family of rhodopsin-type receptors, and suggest that SKR is associated with GTP-binding proteins[17] that effect the second-messenger system of SKR function[3,20].

Because the single cDNA clone encodes a functional SKR, we investigated the specificity of SKR with the oocyte expression system. The three mammalian tachykinins all showed electro-

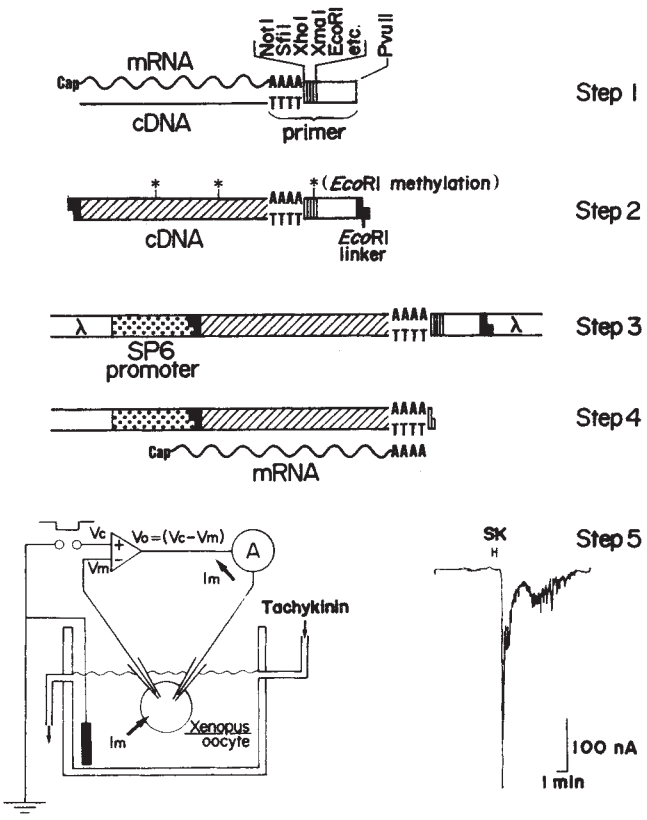


Fig. 1 Strategy for isolating a cDNA clone for bovine SKR. Clone λSKR56 containing the bovine SKR cDNA was isolated as follows. Step 1: To make the primer DNA, the chemically synthesized oligonucleotides containing rare *Not*I and *Sfi*I sites were inserted into the *Sal*I site in the polylinker of pUC19 DNA (ref. 23).

5′-TCGAGCGGCCGCCATGGCCTCGAG-3′
3′-CGCCGGCGGTACCGGAGCTCAGCT-5′

The modified pUC19 DNA (pUC19NS) obtained by transformation was cleaved at the *Pst*I site. About 70 thymidine residues were added to the 3′ end of the *Pst*I site[4] and the DNA fragment containing the poly(dT) tail was isolated after *Pvu*II cleavage. This cDNA primer contained the following restriction sites next to the poly(dT) tail: *Not*I-*Nco*I-*Sfi*I-*Xho*I-*Sal*I-*Xba*I-*Bam*HI-*Xma*I-*Kpn*I-*Sac*I-*Eco*RI. Bovine stomach poly(A) RNA was subjected to sucrose density gradient centrifugation[3], and a fraction giving potent SKR expression in the oocyte expression system was used for cDNA synthesis. Step 2: The single-stranded cDNAs were converted to double-stranded DNAs[24], and *Eco*RI sites were added to both ends of the cDNA sequences through the *Eco*RI linker ligation method[24]. Step 3: The λ vector DNA containing SP6 promoter was constructed as follows. The *Nae*I site of pGEM2 plasmid DNA (Promega Biotec) was converted to the *Bam*HI site by *Bam*HI linker ligation, and the *Bam*HI fragment containing the SP6 promoter was then inserted into the *Bam*HI site of λgt10 after being converted from the single *Eco*RI site of the original λgt10 (ref. 25). The λDNA vector (λgt10SP6) thus constructed contained a single *Eco*RI site immediately downstream of the SP6 promoter, and the cDNA mixture described in step 2 was inserted in this *Eco*RI site. A bovine stomach λcDNA library was then constructed by *in vitro* packaging[4]. Step 4: DNA was extracted from a mixture of $\sim 3\times10^5$ clones derived from the stomach λcDNA library and digested with *Not*I. Digested DNA was then transcribed *in vitro* by SP6 RNA polymerase in the presence of the cap analogue GpppG as described[5]. Step 5: After digestion of template DNA with DNaseI, the mRNAs synthesized *in vitro* were injected into oocytes, and their electrophysiological response to application of substance K was measured by the conventional two-micropipette voltage clamp method[3]. The purification of the cDNA clone for SKR was then carried out by repeating step 4 and step 5 after stepwise fractionations of the response-evoking cDNA mixtures. Finally, a single clone (λSKR56) was obtained. A typical substance K response evoked by the mRNA derived from λSKR56 is also shown: substance K-responsive current (downward deflection indicates inward current) was recorded under voltage clamp at −60 mV. V_c, command potential; V_m, membrane potential; V_o, output of differential amplifier; I_m, current through cell membrane; A, current monitor.

```
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
CTGATCCCTAATGTCTCCTGAACCACATTAGGAGGGACCTGGGTTTGAGTTTATCACCAACTTACTACATGACCTGG  -361
GCTGGTCCCCACTTGGAATGTTTTTCCATTTCTGCAGTGAAGATGCAGGCTGATGATTTAGAAGGATGCCTCTTACTCTGAGACCACGTCTTGGATTCTAGATATCCACGTGACGTTGGA  -241
GAAGGCTGGCCAGGACATGGGAATCTGAAGTGCCTAGGGCTCTGGACCCGGCTGTCTGCTCAGCCTAACTCATCTCTGTTTGAGACTTGGGAGTGACCCCCCAGCTCCCTGGACGCGCCT  -121
GGCTCGGCACCTACTCCTCCTGCTCCCCGCACTCAGAGAGGATTTGTGTGCTCCCACAGGGCACCAGGGCCTGGAGGGAGCCAAGGAGCCTTACCTCTTGTGCCAGACTCAGAGGCGGCC  -1

 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
MetGlyAlaCysValValMetThrAspIleAsnIleSerSerGlyLeuAspSerAsnAlaThrGlyIleThrAlaPheSerMetProGlyTrpGlnLeuAlaLeuTrpThrAlaAlaTyr
ATGGGGGCCTGTGTCGTGATGACCGATATCAACATCTCGTCTGGCCTTGACAGCAATGCCACGGGCATCACAGCCTTCTCCATGCCCGGCTGGCAGCTGGCACTGTGGACCGCAGCCTAC  120

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
LeuAlaLeuValLeuValAlaValMetGlyAsnAlaThrValIleTrpIleIleLeuAlaHisGlnArgMetArgThrValThrAsnTyrPheIleValAsnLeuAlaLeuAlaAspLeu
CTGGCCCTGGTGCTGGTGGCTGTGATGGGCAATGCCACAGTCATCTGGATCATTCTGGCCCATCAGAGGATGCGCACGGTGACCAACTACTTCATCGTCAACCTGGCCCTGGCCGACCTC  240

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
CysMetAlaAlaPheAsnAlaAlaPheAsnPheValTyrAlaSerHisAsnIleTrpTyrPheGlyArgAlaPheCysTyrPheGlnAsnLeuPheProIleThrAlaMetPheValSer
TGCATGGCTGCCTTCAACGCTGCCTTCAACTTCGTCTACGCCAGCCACAACATCTGGTACTTTGGCCGTGCCTTCTGCTACTTCCAAAACCTCTTCCCCATCACAGCCATGTTCGTCAGT  360

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
IleTyrSerMetThrAlaIleAlaAlaAspArgTyrMetAlaIleValHisProPheGlnProArgLeuSerAlaProGlyThrArgAlaValIleAlaGlyIleTrpLeuValAlaLeu
ATCTACTCCATGACTGCCATTGCTGCTGACAGGTACATGGCCATCGTCCACCCCTTCCAGCCCCGGCTCTCAGCTCCTGGCACCAGAGCGGTCATTGCTGGCATCTGGCTGGTGGCCCTG  480

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
AlaLeuAlaPheProGlnCysPheTyrSerThrIleThrThrAspGluGlyAlaThrLysCysValValAlaTrpProGluAspSerGlyGlyLysMetLeuLeuLeuTyrHisLeuIle
GCCCTCGCCTTCCCCCAATGCTTCTATTCCACCATCACCACGGATGAGGGTGCCACCAAGTGCGTGGTGGCCTGGCCTGAAGACAGCGGTGGCAAGATGCTCCTTTTGTACCACCTCATA  600

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
ValIleAlaLeuIleTyrPheLeuProLeuValValMetPheValAlaTyrSerValIleGlyLeuThrLeuTrpArgArgSerValProGlyHisGlnAlaHisGlyAlaAsnLeuArg
GTGATCGCCCTCATTTACTTCCTGCCTCTCGTGGTGATGTTCGTCGCCTACAGTGTCATCGGACTCACGCTCTGGAGACGCTCTGTCCCGGGGCACCAGGCTCATGGTGCCAACTTGCGC  720

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
HisLeuGlnAlaLysLysLysPheValLysThrMetValLeuValValValThrPheAlaIleCysTrpLeuProTyrHisLeuTyrPheIleLeuGlyThrPheGlnGluAspIleTyr
CACCTGCAGGCCAAGAAGAAGTTCGTGAAGACCATGGTGCTGGTGGTGGTGACGTTTGCCATCTGCTGGCTGCCCTACCACCTCTACTTCATCCTGGGCACCTTCCAGGAGGACATCTAC  840

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
CysHisLysPheIleGlnGlnValTyrLeuAlaLeuPheTrpLeuAlaMetSerSerThrMetTyrAsnProIleIleTyrCysCysLeuAsnHisArgPheArgSerGlyPheArgLeu
TGCCACAAGTTCATCCAGCAGGTCTACCTGGCGCTCTTCTGGCTGGCCATGAGCTCCACCATGTACAATCCCATCATTTACTGCTGCCTCAACCACAGGTTTCGCTCTGGATTCCGGCTT  960

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
AlaPheArgCysCysProTrpValThrProThrGluGluAspLysMetGluLeuThrTyrThrProSerLeuSerThrArgValAsnArgCysHisThrLysGluIlePhePheMetSer
GCTTTCCGTTGCTGTCCATGGGTCACACCAACCGAGGAGGACAAGATGGAGCTGACCTACACTCCATCCCTCTCCACGAGGGTCAACAGGTGTCACACTAAAGAGATTTTTTTCATGTCT  1080

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
GlyAspValAlaProSerGluAlaValAsnGlyGlnAlaGluSerProGlnAlaGlyValSerThrGluPro
GGGGATGTGGCCCCCTCTGAGGCTGTCAACGGGCAGGCTGAAAGCCCCCAGGCTGGAGTGTCTACTGAACCCTAAAGCCTAGTGCCTGGCTGCAGGGAGAGGCGTCCGCACCCTTGAGAA  1200

GTAGCCAGTCGAAAGGACCACCTCTGAAAGTTTGGCCCGTGCCTCACCCTTCAGTACATCAATGGAAAGATGAGATGGGGCCAGGAACTCTGCAGTGGGTCCTGCCTCTCCCAGCCAGCC  1320

CCTCAGACAGTGCTTCAGGGGATCCAATAATTCAAAGAAGACAGGGTGAGCTAACACTTATCTAATCCAGGAGTTGCAAGGTCGAATGCCTTCCTGGGGCATACAACATAGAGTAAATTG  1440

AACCCGGTATGAAAGAAAAATTGTAGATATATTGATAAAAGATAACTGACTTCAGTGTCGGGCCAATAGGGAATGGTGGGGACTGGCAAACTGGAGGGCATAGGAATCCAAAGGGGAGAG  1560

TAGCTGATCAGCTCTGGGCAAATGCTGCCATGTGACTGATCCTCTACATCTTTCAAGAGAAGCCAGAAATCCAAATTTGTGAAACTTCCAGAATTGTTAAATTTGGCTCACATTGTTATA  1680

AATGTTGTGCGAACCAAACAAAACACACCAGTGGGCAGTACTCAACCATGACACTGATTTATTATCAGGATAAAGACTCAAGGAAAGAAGCTGCTTAATTCTGCAAAGCCCCTAACCTGT  1800

GAGCCAATCCAGGGCTTCTCTGCTTTAGCATCTAAGAACCCAGGTCCACCCCAGGGGAGATGGTGGTTAAGGGTCAAAAGAAGCCTCAGCTGGGTGGTCTCAAGACAATGACCTTCACAT  1920

TGCCTTCCTCTTTCAGTTGCAAAACTACTTTGCAGGGAGGCTGAGGATGATGGGAGAGATGCCTATAAATTAGCAAGTCCAATACATGTTAAGTGCTTCAT  2021
```

Fig. 2 The cDNA sequence for bovine SKR and its deduced amino-acid sequence. Nucleotide residues are numbered in the 5′ to 3′ direction, beginning with the first residue of the initiation methionine codon of a long open-reading frame of the cDNA sequence, and the nucleotides on the 5′ side of residue 1 are indicated by negative numbers. The predicted amino-acid sequence of SKR is shown above the nucleotide sequence. The cDNA sequence for SKR was determined by the chain termination method after subcloning the cDNA insert of λSKR56 into M13mp18 and M13mp19 (ref. 23).

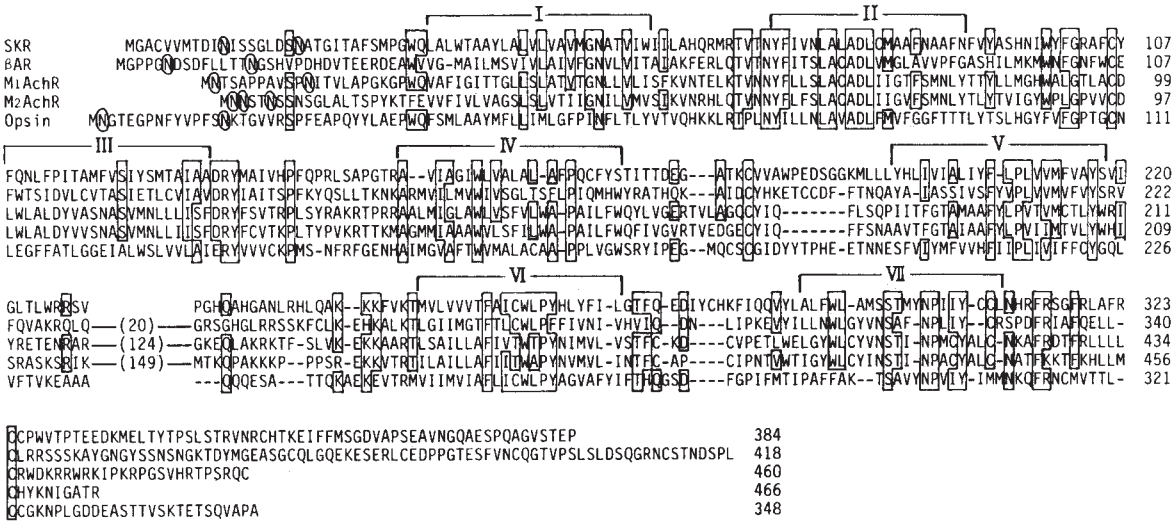
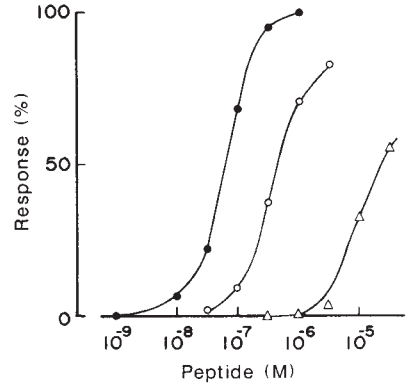


Fig. 3 Alignment of the amino-acid sequences of bovine SKR (top), hamster β_2-adrenergic receptor (βAR, second row), porcine cerebral and cardiac muscarinic acetylcholine receptors (M_1AChR, third row and M_2AChR, fourth row) and bovine opsin (bottom row). (Data for the four latter sequences from refs 11-15). The amino-acid residues enclosed by solid lines represent residues which are identical in at least two of the four reference sequences when compared with the SKR sequence; the non-homologous sequences in loop V-VI are indicated by the sum of residues present in each segment given in parentheses. The degree of sequence identity described in the text was calculated by counting a continuous stretch of gaps as one substitution regardless of its length. Positions of the putative transmembrane segments I-VII of SKR are indicated; the termini of each segment are tentatively assigned on the basis of hydropathicity profile, amino-acid sequence and comparison with rhodopsin-type receptors[13,17]. Circled residues, Asn residues that are potential *N*-glycosylation sites in amino-terminal regions.

Fig. 4 Dose-response curves of substance K (●), neuromedin K (○) and substance P (△) in inducing electrophysiological response of oocytes injected with mRNA derived from cloned SKR cDNA. **Methods.** For experimental convenience, the cDNA insert of λSKR56 was transferred to the plasmid pGEM1 DNA (Promega Biotec); the *Eco*RI fragment containing the whole cDNA sequence together with its following poly(dT-dA) tail and the multiple restriction sites was isolated from λSKR56 and inserted into the *Eco*RI site immediately downstream from the SP6 promoter or the T7 promoter of pGEM1 DNA (pSKR56S and pSKR56T). After digestion of the plasmid DNA with *Not*I, DNA was transcribed by SP6 RNA polymerase (pSKR56S) or T7 RNA polymerase (pSKR56T) in the presence of the capping nucleotide[5]. The mRNA was injected into oocytes (mRNA concentration, 2 ng μl^{-1}; average volume injected per oocyte, ~40 nl). The injected oocytes were incubated at 20 °C for 2-3 days. Electrophysiological measurements were performed at 20-22 °C in Ringer's solution. Responses to tachykinins were recorded under voltage clamp at 0 mV. Each point represents the mean of at least five data.



physiological induction, as indicated in Fig. 4, while no response was induced by the application of the peptides tested that were not tachykinins. The half-maximal concentrations of substance K, neuromedin K and substance P for electrophysiological induction were determined to be 6.0×10^{-8} M, 3.4×10^{-7} M and 8.7×10^{-6} M, respectively. Thus, SKR is the receptor most selective for substance K, but it also has the ability to cross-react with neuromedin K and substance P.

The oocyte system can be used to produce numerous foreign channels as well as their associated neurotransmitter receptors, depending on the source of mRNA injected[21,22]. These include the receptors for amines, amino acids and neuropeptides[3,22]. Furthermore, this system is extremely sensitive; a single SKR clone in the mixture of $\sim3\times10^5$ cDNA clones still produced a clear response to substance K. No convenient biochemical tools for obtaining the purified protein have been available to date for most channels and receptors, but our method will be widely applicable as a direct approach to isolating cDNA clones for these proteins.

We thank K. Takahashi for technical assistance. This work was supported in part by research grants from the Ministry of Education, Science and Culture, Japan.

Received 27 July; accepted 16 September 1987.

1. Quirion, R. *Trends Neurosci.* **8**, 183-185 (1985).
2. Nakanishi, S. *Physiol. Rev.* (in the press).
3. Harada, Y. *et al. J. Neurosci.* (in the press).
4. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
5. Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057-7070 (1984).
6. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
7. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
8. Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
9. Engelman, D. M., Goldman, A. & Steitz, T. A. *Meth. Enzym.* **88**, 81-88 (1982).
10. Ovchinnikov, Y. A. *FEBS Lett.* **148**, 179-191 (1982).
11. Nathans, J. & Hogness, D. S. *Cell* **34**, 807-814 (1983).
12. Dixon, R. A. F. *et al. Nature* **321**, 75-79 (1986).
13. Kubo, T. *et al. Nature* **323**, 411-416 (1986).
14. Kubo, T. *et al. FEBS Lett.* **209**, 367-372 (1986).
15. Peralta, E. G. *et al. Science* **236**, 600-605 (1987).
16. Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28-32 (1975).
17. Dohlman, H. G., Caron, M. G. & Lefkowitz, R. J. *Biochemistry* **26**, 2657-2664 (1987).
18. Hubbard, S. C. & Ivatt, R. J. *A. Rev. Biochem.* **50**, 555-583 (1981).
19. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 852-862 (1975).
20. Buck, S. H. & Burcher, E. *Trends pharmac. Sci.* **7**, 65-68 (1986).
21. Sumikawa, K., Houghton, M., Emtage, J. S., Richards, B. M. & Barnard, E. A. *Nature* **292**, 862-864 (1981).
22. Smart, T. G., Houamed, K. M., van Renterghem, C. & Constanti, A. *Biochem. Soc. Trans.* **15**, 117-122 (1987).
23. Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103-119 (1985).
24. Watson, C. J. & Jackson, J. F. in *DNA Cloning, a Practical Approach* Vol. 1 (ed. Glover, D. M.) 79-88 (IRL, Oxford, 1985).
25. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning, a Practical Approach* Vol. 1 (ed. Glover, D. M.) 49-78 (IRL, Oxford, 1985).

No requirement of cyclic conformation of antagonists in binding to vasopressin receptors

M. Manning*, J. P. Przybylski*, A. Olma*, W. A. Klis*, M. Kruszynski*, N. C. Wo†, G. H. Pelton† & W. H. Sawyer†

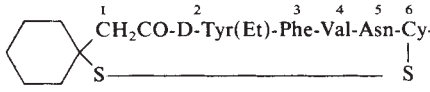
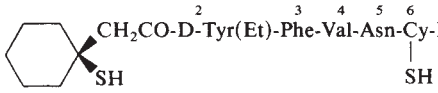
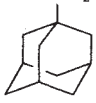
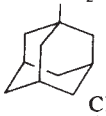
*Department of Biochemistry, Medical College of Ohio, CS 10008, Toledo, Ohio 43699, USA
†Department of Pharmacology, College of Physicians and Surgeons of Columbia University, 630 West 168th Street, New York, New York 10032, USA

Early reports that acyclic analogues of oxytocin and vasopressin (AVP) have drastically reduced agonistic activities established as dogma that an intact hexapeptide ring structure is essential for the pharmacological activities of analogues of neurohypophysial hormones[1]. Thus, virtually all the many hundreds of agonistic and antagonistic analogues of the neurohypophysial peptides that have been reported contain an intact ring[1–3]. Here we report that an intact ring is not essential for binding of antagonistic AVP analogues to vasopressor (V_1) or antidiuretic (V_2) AVP receptors. In fact, one acyclic AVP analogue seems to be about as potent as any previously reported cyclic V_2 antagonist. This finding suggests new possibilities for the design of AVP analogues as pharmacological probes and for therapeutic use. Similar modifications might be useful in the design of analogues of other cyclic peptides, such as calcitonin, somatostatin and the atrial natriuretic factors.

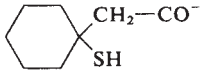
The structures of five acyclic peptides are shown in Table 1. All are derived from the cyclic vasopressin antagonist, [1-(β-mercapto-β,β-cyclopentamethylenepropionic acid), 2-(*O*-ethyl)-D-tyrosine, 4-valine]arginine-vasopressin or d$(CH_2)_5$D-Tyr(Et)VAVP (ref. 4) (Table 1). Peptide 1 is the unoxidized precursor of d$(CH_2)_5$D-Tyr(Et)VAVP containing two sulphydryl groups rather than a disulphide bridge. Peptides 2–5 have α-aminobutyric acid rather than a half-cystine in the 6 position. None of these could, therefore, form a ring involving a disulphide bridge between the 1 and 6 positions. Peptides 3 and 4 have 1-adamantaneacetic acid in the 1 position and peptide 5 has propionic acid in this position. Peptides 3–5 thus contain no sulphydryl groups at all.

The protected intermediates required for the synthesis of peptides 1–5 were synthesized by the solid-phase method[5,6] using modifications previously described for the synthesis of d$(CH_2)_5$D-Tyr(Et)VAVP (ref. 4). Omission of the usual oxidative cyclization with dilute potassium ferricyanide following deblocking[4] gave peptide 1, the uncyclized precursor of d$(CH_2)_5$D-Tyr(Et)VAVP. Peptides 1–5 were shown to be homogeneous by thin-layer chromatography in three solvent

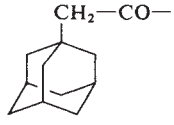
Table 1 Antagonistic potencies of linear analogues of a cyclic vasopressin antagonist

No.	Antagonist*	*Antiantidiuretic* Effective dose† (nmol kg^{-1})	*Antiantidiuretic* pA_2‡	*Antivasopressor* Effective dose† (nmol kg^{-1})	*Antivasopressor* pA_2‡
	 1 2 3 4 5 6 7 8 9 CH_2CO-D-Tyr(Et)-Phe-Val-Asn-Cy-Pro-Arg-Gly-NH_2§	1.1 ± 0.2	7.81 ± 0.07	0.45 ± 0.11	8.22 ± 0.12
1	 2 3 4 5 6 7 8 9 CH_2CO-D-Tyr(Et)-Phe-Val-Asn-Cy-Pro-Arg-Gly-NH_2‖	6.2 ± 1.1	7.05 ± 0.08	2.1 ± 0.2	7.51 ± 0.03
2	CH_2CO-D-Tyr(Et)-Phe-Val-Asn-Abu-Pro-Arg-Gly-NH_2	2.6 ± 0.3	7.42 ± 0.05	1.3 ± 0.2	7.71 ± 0.05
3	 CH_2CO-D-Tyr(Et)-Phe-Val-Asn-Abu-Pro-Arg-Gly-NH_2	2.1 ± 0.1	7.50 ± 0.02	5.7 ± 0.9	7.10 ± 0.07
4	 CH_2CO-D-Tyr(Et)Phe-Val-Asn-Abu-Pro-Arg-Arg-NH_2	0.53 ± 0.07	8.11 ± 0.07	1.2 ± 0.2	7.75 ± 0.07
5	CH_2-CO-D-Tyr(Et)-Phe-Val-Asn-Abu-Pro-Arg-Arg-NH_2 (CH_3)	1.3 ± 0.2	7.74 ± 0.07	14 ± 3	6.71 ± 0.09

* Structures, names, and abbreviations of the substituents in positions 1 and 6:

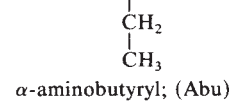


(β-mercapto-β,β)cyclopentamethylenepropoinyl); (d$(CH_2)_5$)



1-adamantaneacetyl-; (Aaa)

CH_3CH_2—CO—, propionyl



α-aminobutyryl; (Abu)

† The effective dose is defined as the dose (in nmol kg^{-1}) of antagonist that reduces the response seen with $2x$ units of agonist to equal the response seen with x units of agonist administered in the absence of antagonist.

‡ Estimated *in vivo* pA_2 values represent the negative logarithms of the effective doses divided by the estimated volume of distribution (67 ml kg^{-1}).

§ [1-(β-mercapto-β,β-cyclopentamethylenepropionic acid), 2-(*O*-ethyl)-D-tyrosine, 4-valine]arginine vasopressin (d$(CH_2)_5$D-Tyr(Et)VAVP): data from ref. 4.

‖ Open chain precursor of d$(CH_2)_5$D-Tyr(Et)VAVP.

systems and gave the expected amino-acid analysis ratios. The V_2 and V_1 antagonistic activities of these peptides were estimated by assays on water-loaded rats anaesthetized with ethanol[7,8] and on phenoxybenzamine-treated rats under urethane anaesthesia[9,10], as elsewhere[3].

All five acyclic analogues antagonize both antidiuretic (V_2) and vasopressor (V_1) responses to vasopressin (Table 1). This clearly demonstrates that an intact ring structure is not essential for recognition and binding of AVP analogues by V_1 or V_2 receptors. Peptide 4 seems about twice as potent a V_2 antagonist as $d(CH_2)_5$D-Tyr(Et)VAVP and is, in fact, one of the most potent V_2 antagonists known[2,3,11-13]. Peptide 5 is about as potent as $d(CH_2)_5$D-Tyr(Et)VAVP, indicating that a bulky hydrophobic grouping in the 1 position, as in peptides 1-4, is not required for effective V_2 antagonism by acyclic AVP analogues.

The surprising discovery that a cyclic conformation is not required for the binding of antagonists to AVP receptors calls for serious reexamination of whether a ring structure is essential for the activation of AVP V_1 and V_2 receptors. Note that a linear fragment of AVP is highly active in delaying extinction of a conditioned avoidance response in rats[14].

These linear antagonists have obvious potential as pharmacological probes of the actions of endogenous AVP (refs 3, 11 and 15). They could also be of value for treating inappropriate secretion of the antidiuretic hormone (SIADH)[16]. Furthermore, these findings provide a new approach to the design of orally active and/or non-peptide antagonists of AVP.

Our discovery of a cyclic peptide antagonist which binds to its specific receptors when linearized raises the possibility that other cyclic peptides such as somatostatin, calcitonin and the atrial natriuretic factors may also not require a cyclic conformation for binding to their specific receptors.

We thank Ms Ann Chlebowski for help with the manuscript. This work was supported by research grants from the National Instititue of General Medical Sciences and the National Institute of Diabetes and Digestive and Kidney diseases.

Note added in proof: It has recently come to our attention that the non-requirement of a ring structure for the elaboration of agonistic activities by salmon calcitonin analogues has already been demonstrated[17].

Received 10 July; accepted 18 September 1987.

1. Berde, B. & Boissonnas, R. A. *in Handbook of Experimental Pharmacology* Vol. 23 (ed. Berde, B.) 802-807 (Springer, Berlin, 1968).
2. Hruby, V. J. & Smith, C. W. in *The Peptides* Vol. 8 (eds Udenfriend, S. & Meienhofer, J.) 77-207 (Academic, Orlando, Florida, 1987).
3. Manning, M., Bankowski, K. & Sawyer, W. H. in *Vasopressin, Principles and Properties* (eds Gash, D. M. & Boer, G. J.) 335-368 (Plenum, New York, 1987)
4. Manning, M. *et al. J. med. Chem.* **25**, 45-50 (1982).
5. Merrifield, R. B. *J. Am. chem. Soc.* **85**, 2149-2154 (1963).
6. Merrifield, R. B. *Science* **232**, 341-347 (1986).
7. Sawyer, W. H. *Endocrinology* **63**, 694-698 (1958).
8. Sawyer, W. H. *et al. Science* **212**, 49-51 (1981).
9. Dekanski, J. *Br. J. Pharmac.* **7**, 567-572 (1952).
10. Dyckes, D. F., Nestor, J. J. Jr, Ferger, M. F. & du Vigneand, V. *J. med. Chem.* **17**, 250-252 (1974).
11. Manning, M. & Sawyer, W. H. *J. cardiovasc. Pharmac.* Suppl. **8**, 529-535 (1986).
12. Ali, F. E. *et al. J. med. Chem.* **29**, 984-988 (1986).
13. Nelson, C. F. *et al. J. med. Chem.* **29**, 2425-2426 (1986).
14. De Wied, D., Gaffori, O., van Ree, J. M. & de Jong, W. *Nature* **308**, 276-278 (1984).
15. Johnston, C. I. *J. Hypertension* **3**, 557-569 (1985).
16. Schrier, R. W. *New Engl. J. Med.* **312**, 1121-1123 (1985).
17. Orlowski, R. C., Epand, R. M. & Stafford, A. R. *Eur. J. Biochem.* **163**, 399-402 (1987).

An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2

Brian Seed

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

Recently the human T cell erythrocyte receptor CD2 has been shown to bind human erythrocytes through LFA-3, a heavily glycosylated surface protein of broad tissue distribution[1,2]. CD2-LFA-3 interactions are important for cytolytic conjugate formation[3,4], for thymocyte adhesion[5], and for T cell activation[6]. A complementary DNA clone encoding LFA-3 was isolated using a novel transient expression system of mouse cells[11]. The cDNA encodes a phospholipid-linked membrane protein whose extracellular domain shares significant homology with CD2. As CD2 is homologous with the neural cell adhesion molecule NCAM in immunoglobulin-like domains[7], cellular adhesion molecules in both neural and lymphoid tissues could have a common ancestor.

Three lymphocyte function-associated structures participate in T cell-target cell adhesion[8,3]; LFA-1 and LFA-2 on the T cell, and LFA-3 on the target cell[9]. LFA-2 is the T cell erythrocyte receptor (CD2) (ref. 9), and recent antibody-blocking studies[4], experiments with purified CD2 and LFA-3 (refs 1, 2, 10), and the results of transfections with cloned CD2 (ref. 11) have shown that LFA-3 is the ligand for CD2. In addition to cytolytic conjugate formation, CD2-LFA-3 interaction mediates thymocyte-thymic epithelium adhesion[5], a reaction which could be important in prothymocyte homing; CD2 is one of the earliest antigens of the T cell lineage, and can be found on rare cells in bone marrow[12]. A sheep homologue of LFA-3 (called T11TS) mediates the binding of sheep erythrocytes to T cells[13], which allows T cell activation when coupled with exposure to a particular antibody recognizing CD2 (ref. 6).

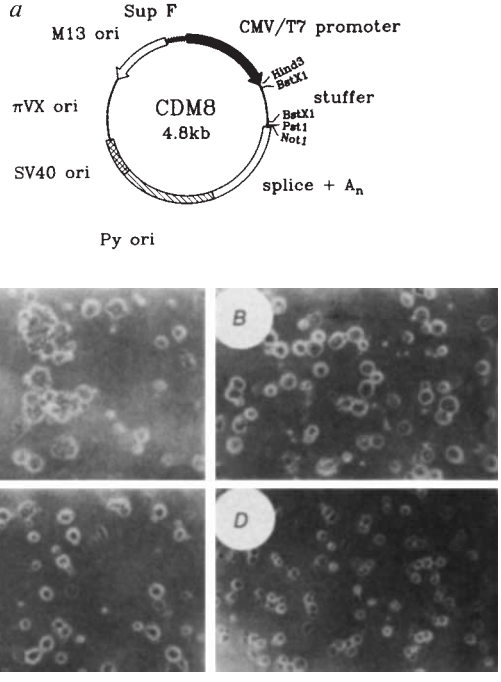


Fig. 1 Isolation of an LFA-3 cDNA clone. *a*, The vector contains both SV40 and polyoma replication origins, a cytomegalovirus/T7 RNA polymerase promoter, and an M13 origin of replication. A library of 1.9×10^6 clones was prepared in CDM8 from 1 μg of poly(A)$^+$ RNA from JY cells and panned as described[11,14]. *b*, Rosette formation between CD2- and LFA-3-transfected COS cells. Cells transfected with cloned DNA were trypsinized 36 h after transfection, transferred to new dishes, collected 12 h later in PBS/0.5 mM EDTA, passed through nylon mesh to eliminate aggregates, and allowed to form rosettes in the presence or absence of anti-CD2 (OKT11) or antibody to LFA-3 as ascites at 1:100. *A*, No antibody; *B*, antibody to CD2; *C*, antibody to LFA-3, *D*, mixture of CD2 and CD1b transfected cells.

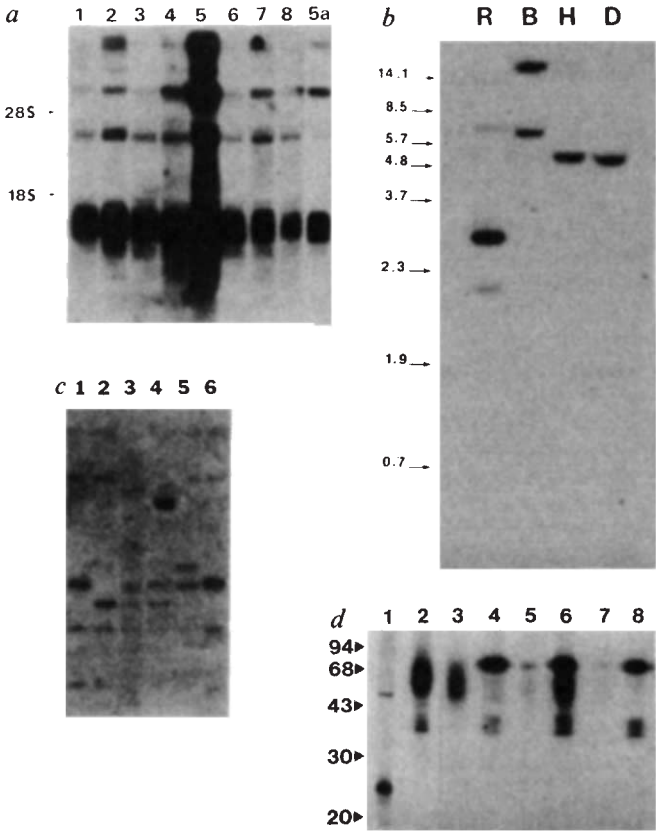


Fig. 2 Characterization of the LFA-3 cDNA clone. *a*, RNA blot hybridization analysis of LFA-3 transcripts. 20 μg total RNA were denatured, electrophoresed through agarose, transferred to nylon membranes, hybridized with LFA-3 cDNA and autoradiographed for either 54 h (lanes 1-8), or 5 h (lane 5a). RNA sources were: 1, HPB-ALL (T cell leukaemia); 2, MOLT-4 (T cell leukaemia); 3, HepG2 (hepatoblastoma); 4, placenta; 5, JY (lymphoblastoid cell line); 6, Raji (Burkitt lymphoma); 7, HL-60 (promeyelocytic leukaemia); and 8, U937 (histiocytic lymphoma). Size of RNA markers is shown to the left. *b*, DNA blot hybridization analysis of LFA-3 genomic DNA. 20 μg placental DNA was restricted, electrophoresed through agarose, transferred to nylon, hybridized and autoradiographed for 36 h. Restriction enzyme digests were: R, *Eco*RI; B, *Bam*HI; H, *Hin*dIII; and D, *Dra*I. DNA size markers (kb) shown to the left. *c*, Arrangement of LFA-3 genomic sequences in DNA from several B-cell lines and placenta. DNA was digested with *Eco*RI and hybridized as for Fig. 2*b*. DNAs were from: lane 1, JY cells; lane 2, Daudi cells; lane 3, CESS cells; lane 4, Namalwa cells; lane 5, peripheral blood lymphocytes; and lane 6, placenta. *d*, Electrophoretic characterization of LFA-3 cDNA protein. Lane 1, *in vitro* transcription/translation product of the LFA-3 insert; 2, LFA-3 immunoprecipitate from JY cells; 3, immunoprecipitate from LFA-3-transfected COS cells; 4, from the supernatant of LFA-3-transfected COS cells; 5, from phospholipase-treated transfected COS cells; 6, from the supernatant of phospholipase-treated cells; 7, from CD4-transfected cells; and 8, from the supernatant of CD4-transfected cells. Cell surface labelling and immunoprecipitations were performed as described[11]. Phospholipase treatment was in 270 mM sucrose/10% PBS at 37 °C for 1 h using 40 μg ml^{-1} of *S. aureus* phosphatidylinositol-specific phospholipase C. M_r markers (*K*) are given on the left.

Human LFA-3 is defined by a single antibody, TS2/9, which reacts with determinants on COS cells. To clone an LFA-3 cDNA by panning techniques[11,14] a new expression vector (CDM8; Fig. 1) was created for the murine WOP3027 (ref. 15) cell-expression system. WOP cells support the replication of plasmids containing polyoma virus origins, and do not react with TS2/9 antibody.

A library was prepared from the lymphoblastoid cell line JY, introduced into WOP cells by spheroplast fusion and subjected to three rounds of panning and reintroduction into *Escherichia coli* as described[11]. Of six plasmids picked after the third round, three tested positive for LFA-3 expression by indirect immunofluorescence in WOP cells. The three positive clones each contained an insert of about 900 base pairs (bp). Expression in COS cells causes a specific rosette formation between CD2-transfected cells and cells transfected with the putative LFA-3 clone (Fig. 1). Rosette formation could be blocked by pretreatment of the cells with either antibody to CD2 or antibody to LFA-3.

RNA blot hybridization analysis (Fig. 2*a*) revealed a major RNA species of about 1.1 kilobases (kb) in length, as well as a number of minor species with similar abundance among RNAs from T cell, B cell, monocytic and hepatic tumour lines and placenta. Ten times more LFA-3 transcript was detected in total RNA prepared from the lymphoblastoid cell line JY. Both the major and minor transcripts were comparably amplified in JY RNA, suggesting a transcription effect. The increased mRNA concentration is consistent with the high degree of LFA-3 expression previously noted in these cells[16].

DNA blot hybridization analysis of LFA-3 genomic DNA gave a simple pattern consistent with a single copy gene (Fig. 2*b*). Examination of the *Eco*RI digest generated from a panel of B-cell and germ-line DNAs showed that the increased expression in the JY line did not result from gene amplification or rearrangement (Fig. 2*c*). A very high degree of fragment length polymorphism was observed about the gene, with three alleles present in six isolates. The comparable spacing of the fragment length suggests that the polymorphism could arise from different numbers of repeat units in a redundant sequence located close to the body of the gene.

In vitro transcription and translation of LFA-3 cDNA gave a band of relative molecular mass (M_r) 26,000 (26K) by SDS-PAGE, and immunoprecipitation of transfected COS cells gave a broad band between 45 and 65K, slightly smaller than that obtained from JY cells (Fig. 2*d*). DNA sequence analysis by the dideoxy method showed that the 874-bp insert contained an open reading frame of 237 residues originating at a methionine codon embedded in the suggested consensus sequence[17] (Fig. 3*a*). The predicted mass of the precursor is 26.7K, in close agreement with the mass of the *in vitro* translation product. A hydrophobic region presumably representing a secretory signal sequence follows the initiation codon; using a likelihood ratio calculation[18], the amino terminus of the mature protein was predicted to occur at Gln31. The amino terminal sequence of the sheep T11TS molecule is 48% homologous to the predicted LFA-3 sequence, and begins at a residue corresponding to Ser29 (T. Hünig, personal communication).

The open reading frame terminates in a series of hydrophobic residues bearing interspersed hydrophilic residues and lacking the basic anchoring residues typically found in surface proteins bearing cytoplasmic domains. This pattern is shared by the carboxy-terminal sequences of proteins attached to the cell by a phosphatidylinositol membrane lipid[19-21]. COS cells transfected with the LFA-3 cDNA were treated with phospholipase specific for phosphatidylinositol, and ~70% of the surface LFA-3 was released (Fig. 2*d*). A similar proportion of phosphatidylinositol-linked alkaline phosphatase has been reported to be released from COS cells transfected with alkaline phosphatase cDNA[22].

The predicted amino-acid sequence was compared to the National Biomedical Research Foundation (NBRF) database, and no significant homologies were found. Further searches among recently published sequences focused on surface antigens known, like LFA-3, to be present on chromosome 1 (ref. 15), to be phospholipid-linked, or involved in T-cell development. Unexpectedly, the CD2 antigen[11,23,24] gave the highest homology score of all proteins, including those in the NBRF database, using the FASTP rapid alignment algorithm. A comparison of

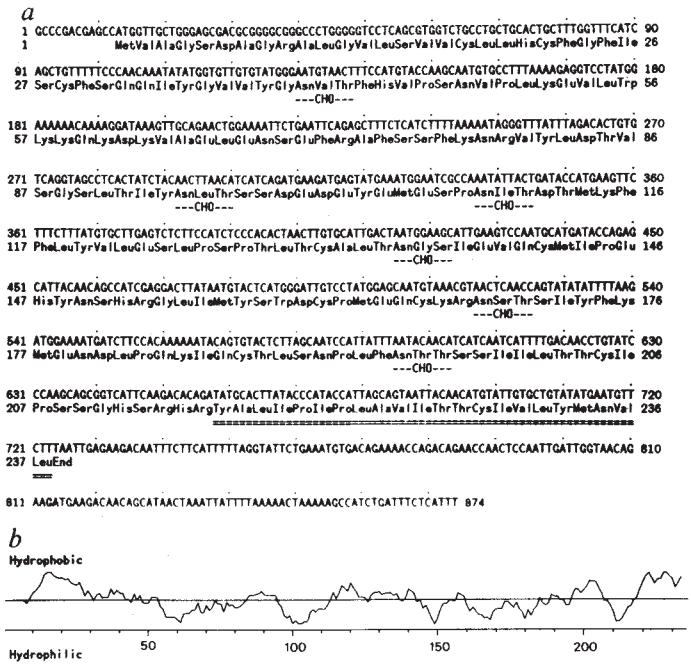


Fig. 3 *a*, Sequence of the LFA-3 cDNA. The sites of potential *N*-linked glycosylation are denoted by the symbol -CHO-; the hydrophobic carboxyl terminus is underscored. *b*, Hydropathicity profile of the amino-acid sequence in *a*.

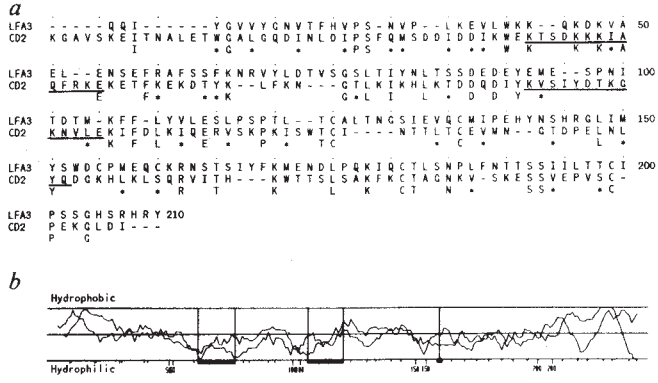


Fig. 4 *a*, Optimal alignment of the extracellular portions of LFA-3 and CD2 by the ALIGN program of the Protein Identification Resource (NBRF) (ref. 25). Conserved residues are displayed beneath the aligned sequences; asterisks were displayed if the residues were closely related. Epitope regions of CD2 identified elsewhere[28] are shown underlined. *b*, Superposition of the hydropathicity plots of LFA-3 and CD2. The amino-terminal sequences of both proteins were compared through the carboxy-terminal hydrophobic sequences. Solid bars denote the antigenic regions identified in CD2.

LFA-3 and CD2 using the ALIGN program of the NBRF sequence comparison package gave optimal alignment of the extracellular domains, as shown in Fig. 4. Monte Carlo simulation of the alignment of 500 randomly permuted variants of the two sequences gave a mean score 5.2 s.d. lower than the alignment score computed for LFA-3 and CD2, which corresponds to a probability of $\sim 10^{-7}$ for spontaneous occurrence of an equally good or better match between two proteins of identical composition[25]. As the homology extends throughout the external domain of the two molecules, the proteins could be distantly related. An alternative explanation, that convergent selective pressures have shaped essentially similar molecules from dissimilar archetypes, requires that multiple structural features of the two molecules be selected. Alignment of the hydropathicity profiles (Fig. 4) shows that, despite substantial divergence, the two proteins have strikingly similar gross organization. A precedent for the hypothetical ancestral progenitor may be found in the homotypic neural cell adhesion molecule NCAM, which adopts both phosphatidylinositol-linked and conventional transmembrane forms[26,27]. Moreover, CD2 is significantly homologous to two NCAM segments of ~200 residues which span domains II and III, and IV and V (ref. 7 and A. F. Williams, personal communication). Thus the heterotypic lymphoid and homotypic neural cell adhesion reactions could share a common evolutionary origin.

I thank Jen Sheen for assistance with the *in vitro* transcription and translation, Steven Herrmann for phospholipase C, Ivan Stamenkovic for the B-cell filter and B-cell and germ line DNA, Alan Williams for discussion and communication of the NCAM/CD2 homology, Steven Burakoff for TS2/9 antibody and encouragement, Thomas Hünig for the T11TS sequence, and David Simmons, Andrew Peterson, the referees, and members of the group for criticism. This work was supported by a grant from Hoechst AG.

Note added in proof: A cDNA encoding a transmembrane form of LFA-3 has recently been isolated (B. Wallner *et al. J. exp. Med.*, in the press).

Received 8 June; accepted 22 September 1987.

1. Plunkett, M. L., Sanders, M. E., Selvaraj, P., Dustin, M. L. & Springer, T. A. *J. exp. Med.* **165**, 664–676 (1987).
2. Selvaraj, P. *et al. Nature* **326**, 400–403 (1987).
3. Krensky, A. M., Robbins, E., Springer, T. A. & Burakoff, S. J. *J. Immun.* **132**, 2180–2182 (1984).
4. Shaw, S. *et al. Nature* **323**, 262–264 (1986).
5. Vollger, L. W., Tuck, D. T., Springer, T. A., Haynes, B. F. & Singer, K. H. *J. Immun.* **138**, 358–363 (1987).
6. Hünig, T. R., Tiefenthaler, G., Meyer zum Büschenfelde, K.-H. & Meuer, S. C. *Nature* **326**, 298–301 (1987).
7. Williams, A. F. *Immun. Today* **8**, 298–303 (1987).
8. Sanchez-Madrid, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7489–7493 (1982).
9. Krensky, A. M. *et al. J. Immun.* **131**, 611–616 (1983).
10. Dustin, M. L., Sanders, M. E., Shaw, S. & Springer, T. A. *J. exp. Med.* **165**, 677–692 (1987).
11. Seed, B. & Aruffo, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3365–3369 (1987).
12. van Dongen, J. J. M. *et al. J. Immun.* **135**, 3144–3150 (1985).
13. Hünig, T. *J. exp. Med.* **162**, 890–901 (1985).
14. Aruffo, A. & Seed, B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
15. Dailey, L. & Basilico, C. *J. Virol.* **54**, 739–749 (1985).
16. Barbosa, J. A. *et al. J. Immun.* **136**, 3085–3091 (1986).
17. Kozak, M. *Microbiol. Rev.* **47**, 1–45 (1983).
18. von Heijne, G. *Nucleic Acids Res.* **14**, 4683–4690 (1986).
19. Low, M. G. & Kincade, P. A. *Nature* **318**, 62–64 (1985).
20. Tse, A. G. D., Barclay, A. N., Watts, A. & Williams, A. F. *Science* **230**, 1003–1008 (1985).
21. Low, M. G., Ferguson, M. A. J., Futerman, A. H. & Silman, I. *Trends biochem. Sci.* **11**, 212–215 (1986).
22. Berger, J., Howard, A. D., Gerber, L., Cullen, B. R. & Udenfried, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4885–4889 (1987).
23. Sewell, W. A., Brown, M. H., Dunne, J., Owen, M. J. & Crumpton, M. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8718–8722 (1986).
24. Sayre, P. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 2941–2945 (1987).
25. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. *Meth. Enzym.* **91**, 524–545 (1983).
26. He, H. T., Barbet, J., Chaix, J. C. & Goridis, C. *EMBO J.* **5**, 2489–2494 (1986).
27. Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenbury, R. & Edelman, G. M. *Science* **236**, 799–806 (1987).
28. Peterson, A. & Seed, B. *Nature* **329**, 842–846 (1982).

Monoclonal antibody and ligand binding sites of the T cell erythrocyte receptor (CD2)

Andrew Peterson & Brian Seed

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

The human T cell erythrocyte receptor (CD2 antigen) allows thymocytes and mature T cells to adhere to thymic epithelium and target cells through a cell surface protein, LFA-3 (refs 1–6). Monoclonal antibodies recognizing CD2 can either block adhesion or, in certain combinations, induce an antigen-independent T cell activation[7–9]. We have identified the binding sites for 16 monoclonal antibodies against CD2 by a rapid and generally applicable

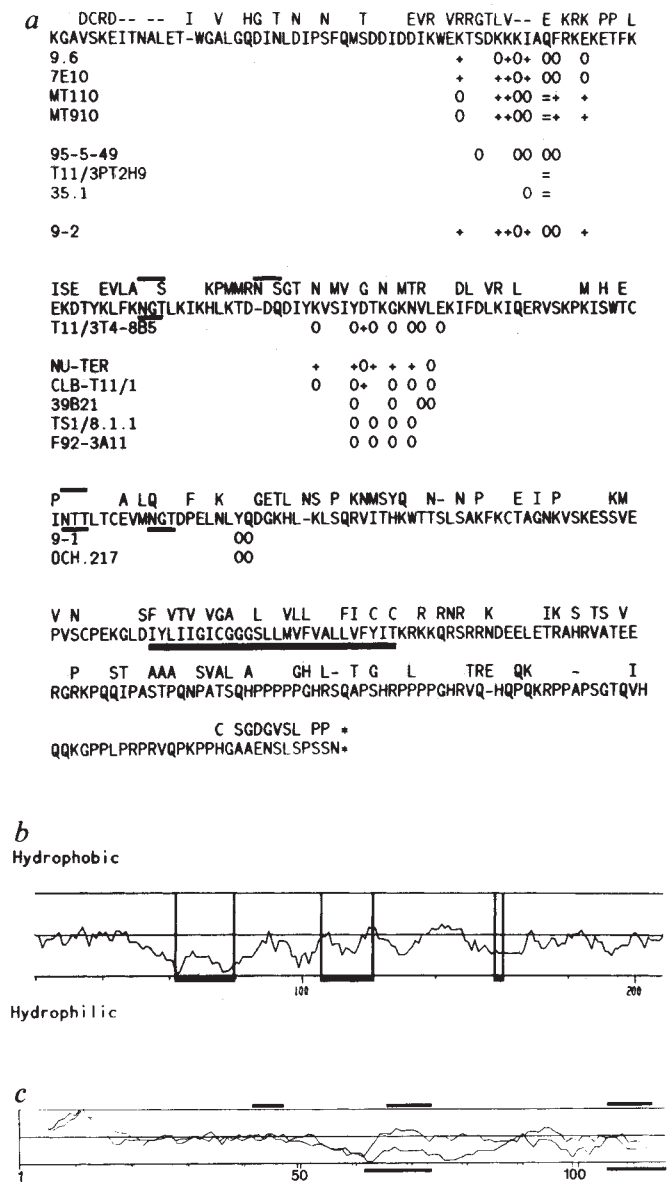


Fig. 1 *a*, The primary amino-acid sequence of the human CD2 protein[6,24,25]. The mouse sequence is shown above the human sequence only where it differs from the human sequence[26]. The large bar indicates the extent of the transmembrane domain. The smaller bars above the amino-acid sequence indicate potential *N*-linked glycosylation sites in the mouse sequence; those below indicate potential sites in the human sequence. The antibodies used are shown along the left margin. The 0 symbol under the primary sequence indicates either that a mutant has a substitution at that position or that indirect immunofluorescence of a mutant obtained with another antibody showed loss of reactivity; + indicates retention of reactivity for all variants examined and = indicates that only a proline substitution at that position affects reactivity *b*, Hydropathicity profile of the first 190 amino acids of CD2 (the extracellular domain) showing the location of the three epitopic regions. *c*, Superposition of the hydropathicity profiles of the first 115 residues of CD2 and the human immunoglobulin κ variable region, Vh (V-III) (ref. 31). κ V-region hypervariable sequences are shown as black bars above the profile, and CD2 ligand binding domains as black bars below the profile. Alignment of the domains shown gives an ALIGN score[32] of 3.7 s.d. above the mean, corresponding to a probability of $\approx 10^{-4}$ for spontaneous occurrence of an equally good or better match.

Methods. The 600 nucleotides of CD2 sequence following position 63 of ref. 6 were synthesized in a collection of twenty 33-mer oligonucleotides, each overlapping its predecessor by three bases. The monomer reagents for the synthesis contained 95% of the wild-type base and 5% of a mixtures of the other 3 bases at each position. A pool of mutants was obtained from each mutagenized 33-mer by oligonucleotide directed mutagenesis of πH3MCD2 (ref. 27) as described[28,29] except that AMV reverse transcriptase was used instead of T4 DNA polymerase. Mutants were selected following spheroplast fusion into COS cells. 48 h post-fusion the COS cells were removed from the culture dish using PBS containing 5 mM EDTA. Antibody incubations and washes were performed as described[6]. The cells were incubated with 0.1% by volume of the negative selection antibody, washed, incubated with 5 μg ml^{-1} of rabbit anti-mouse immunoglobulin antibody (Rockland), washed and incubated for 30 min at 37 °C in 2 mls of 50% rabbit complement (Pel-Freez), 50% Dulbecco's media (GIBCO). After complement lysis the cells were washed, incubated with the positive selection antibody, washed and added to goat anti-mouse immunoglobulin coated dishes as described[6]. Cells adhering to the dish were lysed and the recovered plasmid DNA was transformed into *E. coli*. Mutants were identified by DEAE dextran transfection of COS cells in a 35 mm well with 20% of the plasmid DNA from a 1.5 ml miniprep. The cells were assayed sequentially for binding of the negative and positive selection antibody 48 h post transfection by indirect immunofluorescence. Mutants were sequenced using the chain termination method[30]. In all cases the mutations fell within the span of a single oligonucleotide.

mutational analysis. The binding sites fall in three discrete regions: antibodies that participate in activation and block erythrocyte adhesion bind to the first region; antibodies that block adhesion bind to the second region; and antibodies that participate in activation but do not block adhesion bind to the third region. A large number of mutations selected for loss of antibody reactivity in the first two regions also weaken the CD2–LFA-3 interaction. Good agreement was observed between mutational lesions blocking LFA-3 binding and lesions blocking binding by activating antibodies, which supports the view that such antibodies induce T cell activation by mimicking the effect of LFA-3 binding. CD2 sequences that participate in LFA-3 binding correspond to immunoglobulin variable region hypervariable sequences when the homologous domains are aligned.

To isolate epitope loss mutants, COS cells were transfected with a pool of mutagenized plasmids, cultured for 48 hours, collected and sequentially treated with an anti-CD2 monoclonal antibody, rabbit anti-mouse immunoglobulin antibody, and complement. Because spontaneous deletion mutants arise frequently in COS cells[10,11], a positive selection step was included; the cells spared by complement treatment were treated with antibody recognizing a distinct CD2 epitope(s) and allowed to adhere to dishes coated with goat anti-mouse immunoglobulin antibody[12]. Plasmid DNA recovered from the adherent cells[13] was transformed into *Escherichia coli*, amplified, and reintroduced into COS cells for further rounds as appropriate. At the end of the selection process DNA from individual bacterial colonies was transfected into COS cells which were then scored for antibody binding. The antibodies used for mutant isolation are shown in Table 1. The results of the mutant selections are summarized in Figs 1 and 2.

The mutants are described below by a wild-type residue/mutant residue convention, so that Lys-48Asn, for example, means that the lysine at position 48 has been replaced with an asparagine. 114 Primary mutants were isolated, resulting in a collection of 47 different amino-acid sequence variants. The variation falls in three discrete regions. Region 1 is centred about Lys48 and contains mutations for the antibodies (9.6, 7E10, MT110 and MT910; group I antibodies) which, together with mAb 9-1, can induce IL-2 synthesis in T cells (B. Bierer and A.P., unpublished observations and ref. 14). All but one (9-2) of the other antibodies giving mutations in region 1 have been reported to induce IL-2 receptors but not IL-2 in collaboration with mAb 9-1 (ref. 14). Region 2 is centred about Gly95. Most of the antibodies recognizing region 2 have little effect on T cell activation when used with mAb 9-1. Region 3 is represented by a single mutation which causes loss of reactivity with both 9-1 and OCH217.

The ability of the mutant CD2 proteins to promote adhesion of human erythrocytes to transfected COS cells mediated by

a

42	43	44	45	46	47	48	49	51	52	53	54	55	56	57	- Mab	+ Mab	rosette	MFI
Lys	Thr	Ser	Asp	Lys	Lys	Lys	Ile	Ala	Gln	Phe	Arg	Lys	Glu	Lys				
					Met										9.6	35.1	-	192
					Asn										9.6	35.1	-	159
					Glu										9.6	35.1	-	182
									Arg						9.6	35.1	+	169
										Val			Gly		9.6	35.1	+/-	150
									Pro						35.1	all 16	-	170
					Glu										MT910	all 16		
									Pro						MT910	all 16		
Arg							Asn								MT110	all 16	-	145
Asn									Thr						95-5-49	all 16	-	198
					Asn										95-5-49	all 16		
					Glu										95-5-49	all 16		
									Arg						9-2	all 16		
							Pro								7E10	all 16	-	176
					Thr				Glu						7E10	all 16	-	196
					Asn		Glu								7E10	all 16	-	187
									Pro						7E10	all 16		
					Asn				Pro						7E10	all 16	-	167
									Leu						9.6	35.1	-	239
			Asn		Asn										9.6	35.1	-	176
					Glu					Ile					9.6	35.1	-	160
			Met		Asn				His						9.6	35.1	-	158
									ArgSer						9.6	35.1	+/-	167
					Met				Gln						9.6	35.1		
					Met				Arg						9.6	35.1	-	156
			Asn		Asn					Leu					9.6	35.1	-	160
					Pro									9.6	35.1			
					Glu										9-2	35.1		
									Leu						9-2	35.1	NT	
					Gln										9-2	35.1	-	176
									ArgLeu						9-2	35.1	+/-	160
					Glu					Ile					9-2	35.1		
					Met				Arg						9-2	35.1		
					Glu					Ser					9-2	35.1	-	194
									Glu						9-2	35.1	-	194
					Met										9-2	35.1		
					Thr										9-2	35.1	-	158
					Met				Arg						9-2	35.1		
					GlnVal										9-2	35.1	-	196
					Glu					Ile					7E10	35.1		
										Leu					7E10	35.1		
					Glu					Ile					7E10	35.1		
			Arg						Lys						7E10	35.1	+/-	183
							Glu								7E10	35.1	-	188
									ArgLeu						7E10	35.1		
									Arg						7E10	35.1		
						Gln									35.1	7 others	-	137

b

87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	- Mab	+ Mab	rosette	MFI
Lys	Val	Ser	Ile	Tyr	Asp	Thr	Lys	Gly	Lys	Asn	Val	Leu	Glu	Lys				
				His											NU-TER	all 16	+/-*	245
								Glu				Ser			T11/3T4-8B5	all 16	-	215
					Ser		Val		Ile						CLB-T11/1	all 16	-	243
				Asp											39B21	all 16	+/-*	231
				Val											NU-TER	all 16	-	239
				Ala											NU-TER	all 16	-	230
				Tyr								Phe			NU-TER	all 16	-	216
												Ser			NU-TER	all 16	-	245
				Cys											T11/3T4-8B5	all 16	-	242
Asn															T11/3T4-8B5	all 16	-	241
											Ile		Lys		T11/3T4-8B5	all 16	+/-	225
								Glu							T11/3T4-8B5	all 16	-	177
										Lys			Lys		T11/3T4-8B5	all 16	-	202
				Asp											T11/3T4-8B5	all 16		
						Ser		Val		Ile					T11/3T4-8B5	all 16		
Asn															CLB-T11/1	all 16		
								Glu		Lys		Ser			CLB-T11/1	all 16	-	226
				Cys											CLB-T11/1	all 16		
												Ser			CLB-T11/1	all 16		
								Glu							CLB-T11/1	all 16		
								Glu				Ser			CLB-T11/1	all 16		
												Ser			39B21	all 16		
											Val	Ser			39B21	all 16		
								Glu				Ser			39B21	all 16		
					Ser		Val		Ile						TS1/8.1.1	all 16		
				Asp											TS1/8.1.1	all 16		
				Cys											TS1/8.1.1	all 16		

d

Fig. 2 *a*, The mutant collection defining epitope region 1. CD2 residues 42–57 are shown above the amino-acid substitution encoded by each mutant. The first column on the right shows the antibody used for negative selection, the second column shows the positive selection antibody(s). 'All 16' indicates that all 16 monoclonals in Table 1 were combined and used for the positive-selection step. '7 others' indicates that the 7 antibodies recognizing region 1 were combined and used for the positive-selection step. The third column shows the erythrocyte rosetting phenotype of the mutant: + wild type, +/− partial rosetting, − no detectable rosetting (see Fig. 3). A blank denotes a variant shown elsewhere in the table. *These mutants rosetted much more poorly than the others scored as +/−. The column on the far right shows the mean fluorescence intensity MFI measured by flow cytometry of COS cells expressing the mutant CD2s. *b*, The mutant collection defining epitope region 2. CD2 residues 86–101 are shown above the mutant substitutions. Other notations are as in *a*.

Methods. Cells stained by indirect immunofluorescence were judged to express antigen if ≥3% of the transfected population had an MFI ≥40. Only 0.1% of mock-transfected cells had an MFI ≥ 40, and the mean MFI of the remainder was 4. COS cells transfected with wild-type CD2 gave MFI values of 179 for mAb Nu-Ter (used for region 1 mutants) and 239 for mAb 35.1 (used for region 2 mutants). The mutants listed directly under the primary sequence were from a pool of plasmids mutagenized by oligonucleotides spanning the extracellular domain.

The mutants listed under the bar were obtained using plasmids mutagenized by oligonucleotides encoding the span of the bar.

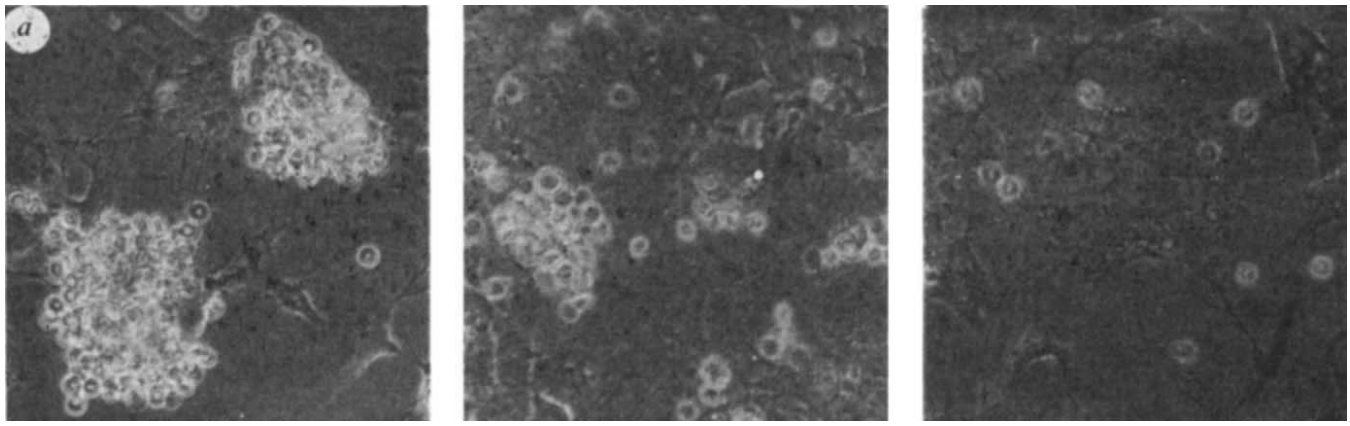
LFA-3 was measured by a qualitative erthyrocyte rosette assay. Three phenotypes were scored: wild-type, partial, and non-rosetting, as illustrated in Fig. 3 and summarized in Fig. 2. Many of the mutations leading to changes in regions 1 and 2 dramatically reduced rosetting. To examine this further, a few mutants were created by specific oligonucleotide mutagenesis. Substitution of asparagine or alanine for lysine at each of positions 46, 47, and 48 demonstrated a striking correlation between the binding of antibody mAb 9.6 and erythrocyte adhesion; Lys46Asn/Ala showed a modest effect on both Mab 9.6 and erthrocyte binding, Lys47Asn/Ala had no effect on either, and Lys48Asn/Ala completely abolished both (Fig. 3*b*). Similarly, residue 51 was important for both erthrocyte and 9.6 binding, whereas residue 52 had only a weak effect on each.

Using different antibodies and mutants, we have also shown that Lys48 is important in the interaction of CD2 with group I antibodies and with LFA-3 (Figs 1, 2 and 3). For example, the mutant Lys48Glu is unreactive with all the group I antibodies, and none of the molecules substituted at Lys48 has any detectable rosetting activity. The behaviour of molecules with substitutions at Lys48 supports the idea that group I antibodies mimic the effect of LFA-3 binding in provoking T cell proliferation.

Although some residues can directly determine antibody reactivity, others of secondary importance for antibody affinity can be identified, because they are frequently altered in association with other changes. For example, a Lys46Asn substitution is frequently found in mutants which do not bind mAb 9.6, but by itself it has little effect on antibody binding. A similar phenomenon may be present in the repeated isolation of position 51, 52 double mutants (Fig. 2).

Amino-acid substitutions could lead to loss of antibody or LFA-3 binding either by elimination of a specific interaction or by causing a local denaturation. Some patterns of antibody binding or rosetting argue against this latter possibility. For example, all the molecules substituted at Lys48 still bind mAb 35.1, which is sensitive to changes at Ile49. Similarly, neither mAb 9.6 or LFA-3 can bind the Gln51Leu variant, which is recognized by antibodies 7E10 and 9-2. A Gln51Arg variant is unreactive with 7E10 and 9-2, but rosettes erthrocytes in the same way as the wild type. In the second epitopic region, Tyr91Asp causes loss of rosetting, but antibody Nu-Ter binding is unaffected, even though many substitutions at position 92 eliminate Nu-Ter reactivity.

However, a Gln51Pro substitution may induce a local denaturation, since proline residues restrict alpha helix formation[15] and none of the antibodies recognizing the first region react with



b Erythrocyte rosetting

		+/-	+	-	-			-	+/-	
Ser	Asp	Lys	Lys	Lys	Ile	Ala	Gln	Phe	Arg	Lys
		+/-	+	-	+			-	+/-	

Antibody 9.6 binding

Fig. 3 *a*, Adhesion of human erthrocytes to transfected COS cells. The left photograph shows wild-type rosettes, the centre shows partial rosettes and the right photograph no rosettes. Wild-type rosettes completely obscure the transfected COS cells, and are macroscopically visible. Partial rosettes are microscopically visible and leave some transfected COS cells exposed. Absence of rosetting was scored if the mutant was indistinguisable from a negative control (CD8 expressing COS cells), that is, no rosettes were found after careful scanning of the plate. *b*, sensitivity of erythrocyte rosetting and 9.6 binding to changes at specific amino-acid positions. + indicates that rosetting or binding is retained upon substitution at that position. +/− indicates that substitution has some effect, namely a partial rosetting phenotype or minimal effect on antibody binding. − indicates that rosetting or antibody binding is eliminated by a single amino-acid substitution at that position.

Gln51Pro. mAbs 35.1 and T11/3PT2H9 gave Gln51Pro exclusively when all 16 antibodies were used for positive selection. Because frequent isolation of Gln51Pro was observed with other mAbs, many of the mutants in the first epitopic region (Figs 1 and 2) were isolated using mAb 35.1 as the only positive selection antibody.

To isolate a 35.1$^-$ mutant other than Gln51Pro, only the antibodies failing to bind to this variant were used for positive selection. Three cycles of enrichment gave a single 35.1$^-$ Ile49Gln mutant altered in all three bases of the original codon. This unusual mutation suggests that the affinity of 35.1 antibody derives from multiple conformational features of CD2, so that substitution for a single feature only rarely greatly decreases affinity. The Gln51Pro mutation may eliminate several of these interactions by gross alteration of the local secondary structure. Because the affinity of the 35.1 antibody is comparable to that of antibody 9.6 (ref. 16), the unusual mutational pattern of this antibody probably arises from a different type of binding and not simply from a stronger interaction.

Table 1 Antibodies

Antibody	Isotype
9.6	IgG_{2a}
7E10	IgG_{2b}
MT910	IgG_1
MT110	IgG_1
95-5-49	?
35.1	IgG_{2a}
T11/3PT2H9	IgG_1
T11/3T4-8B5	IgG_{2a}
9-2	IgM
Nu-Ter	IgG_1
CLB-T11/1	IgG_1
39B21	IgG^*_{2a}
TS1/8.1.1	IgG_1
F92-3A11	IgG_1
9-1	IgG_3
OCH217	IgM

A partial panel of anti-CD2 monoclonal antibodies was obtained. The first four antibodies (9.6, 7E10, MT910, MT110) can each induce IL-2 release from T cells expressing CD2 in the presence of antibody 9-1. A more complete functional analysis of the antibodies can be found in ref. 14.

* Antibody 39B21 is a rat monoclonal and all others are mouse antibodies.

Only one mutant was found with the two antibodies recognizing region 3, a Tyr140Asn and Gln141His double substitution. Both of these antibodies, however, react only weakly with the CD2 molecule expressed on COS cells, which compares with their weak reactivity with CD2 on unactivated T cells[14]. Previous activation of T cells or incubation with a group I antibody is necessary to make the 9-1 epitope available[14]. The rapid acquisition of mAb 9-1 reactivity suggests that it is caused by a conformational change in the molecule and not by *de novo* synthesis of a different species[17].

To further study the interaction of antibodies with each of the two major antibody-binding regions, a large number of mutants were isolated using a CD2 preparation mutagenized by only one or a few oligonucleotides (Fig. 2; see also Fig. 1 legend). Mutants were obtained from such plasmid pools at a frequency of 75–100% after a single round of selection. This allowed a large number of amino-acid variants to be quickly isolated. In the first epitope region the antibodies 7E10, 9-2 and 9.6 were chosen for intensive study because they appear to contact many of the same amino-acid residues (Figs 1 and 2). Two of the antibodies can function, together with mAb 9-1, in T cell activation but the third (9-2) cannot[14]. Each antibody gave rise to a slightly different range of mutations (Fig. 2): the 9-2$^-$ mutations span only 5 residues compared to 8 for the 7E10$^-$ mutations and 10 for the 9.6$^-$ mutations. 9-2 is the only IgM antibody which recognizes region 1, and its inability to activate could be due to a decreased affinity, or to steric interference with 9-1. A large number of mutants were similarly isolated in the second epitopic region (Fig. 2).

It has been proposed that CD2 mediates both cell-cell adhesion and antigen-independent activation reactions. The former function is well-established[2,3,18], but the case for the latter still rests on the unique properties of antibodies such as 9-1 in triggering proliferation in the presence of either group I antibodies[7,14] or sheep erthrocytes[19-21]. The first epitope region we have identified is probably important in both the adhesion and activation functions of CD2. Binding of the first region by antibodies allows CD2 to respond to subsequent binding of antibody 9-1 and we predict that LFA-3 binding to CD2 would allow comparable response to 9-1. If so, the adhesion and activation mediated by CD2 are intimately related and not distinct functions.

Because 9-1 does not block erythrocyte rosetting, and because a CD2 variant which does not react with mAb 9-1 still binds erythrocytes, it is unlikely that LFA-3 binding alone can cause activation; further analysis of the region recognized by 9-1 antibodies is necessary for insight into the activation mechanism.

Recently the case for inclusion of CD2 in the immunoglobulin superfamily[22] has been strengthened by discovery of highly significant homologies between CD2 and non-immunoglobulin members of the family[23]. Alignment of the N-terminal 115 residues of CD2 with immunoglobulin κ variable sequences shows that the CD2 regions 1 and 2 correspond to the locations of light chain hypervariable (antibody-combining site) regions 2 and 3 (Fig. 1). This suggests that CD2 ligand-binding sites are phylogenetically related to variable region-combining sites, and supports the idea that adhesion interactions between members of the immunoglobulin superfamily can mimic antibody-antigen interaction.

We thank Barbara Bierer, Steve Burakoff and Sandro Aruffo for discussion, Mary Honma and Dave Simmons for comments on the manuscript, Bill Sewell and Mike Crumpton for a preprint of their work, Neil Clipstone for a sequence reading, Lerzan Kizilay and John Smith for oligonucleotides and Ivan Stamenkovic for discussion and blood sampling. We thank the organizers and participants of the Third International Workshop on Human Leukocyte Differentiation Antigens for making antibodies available. This work was supported by a grant from Hoechst AG.

Received 8 June; accepted 22 September 1987.

1. Vollger, L. W., Tuck, D. T., Springer, T. A., Haynes, B. F. & Singer, K. H. *J. Immun.* **138**, 358-363 (1987).
2. Shaw, S. *et al. Nature* **323**, 262-264 (1986).
3. Krensky, A. M., Robbins, E., Springer, T. A. & Burakoff, S. J. *J. Immun.* **132**, 2180-2182 (1984).
4. Dustin, M. L., Sanders, M. E., Shaw, S. & Springer, T. A. *J. exp. Med.* **165**, 677-692 (1987).
5. Selvaraj, P. *et al. Nature* **326**, 400-403 (1987).
6. Seed, B. & Aruffo, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3365-3369 (1987).
7. Meuer, S. C. *et al. Cell* **36**, 897-906 (1984).
8. Olive, D. *et al. Eur. J. Immun.* **16**, 1063-1068 (1986).
9. Huet, S. *et al. J. Immun.* **137**, 1420-1424 (1986).
10. Calos, M. P., Lebkowski, J. S. & Botchan, M. R. *Proc. natn. Sci. U.S.A.* **80**, 3015-3019 (1983).
11. Razzaque, A., Mizusawa, H. & Seidman, M. M. *Proc. natn. Sci. U.S.A.* **80**, 3010-3014 (1983).
12. Wysocki, L. J. & Sato, V. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2844-2848 (1978).
13. Hirt, B. *J. molec. Biol.* **26**, 365-369 (1967).
14. Yang, S. Y. Rhee, S., Angelos, G. & DuPoint, B in *Leukocyte Typing, III* (ed. McMichael, A. J.) 113-116 (OxfordUniversity Press, 1987).
15. Chou, P. Y. & Fasman, G. D. *A. Rev. Biochem.* **47**, 251-276 (1978).
16. Martin, P. J. *et al. J. Immun.* **131**, 180-185 (1983).
17. Yang, S. Y., Chuiab, S. & Dupoint, B. *J. Immun.* **137**, 1097-1100 (1986).
18. Sanchez-Madrid, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7489-7493 (1982).
19. Hünig, T. R., Tiefenthuler, G., Meyer zum Büschenfelde, K.-H. & Meuer, S. C. *Nature* **326**, 298-301 (1987).
20. Hünig, T. *J. exp. Med.* **162**, 890-901 (1985).
21. Hünig, T. R. *J. Immun.* **136**, 2103-2108 (1986).
22. Williams, A. F., Barclay, A. N., Clark, S. J., Paterson, D. J. & Willis, A. C. *J. exp. Med.* **165**, 368-380 (1987).
23. Williams, A. F. *Immun. Today* **8**, 298-303 (1987).
24. Sewell, W. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 8718-8722 (1986).
25. Sayre, P. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 2941-2945 (1987).
26. Sewell, W. A. *et al. Eur. J. Immun.* **17**, 1015-1020 (1987).
27. Aruffo, A. & Seed, B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
28. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488-492 (1985).
29. Levinson, A., Silver, D. & Seed, B. *J. molec. appl. Genet.* **2**, 507-517 (1984).
30. Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
31. Pech, M. & Zachau, H. G. *Nucleic Acids Res.* **12**, 9229-9236 (1984).
32. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. *Meth. Enzym.* **91**, 524-545 (1983).

Anchoring mechanisms for LFA-3 cell adhesion glycoprotein at membrane surface

Michael L. Dustin, Periasamy Selvaraj, Robert J. Mattaliano* & Timothy A. Springer

Laboratory of Membrane Immunochemistry, Dana Farber Cancer Institute, and Harvard Medical School, Boston Massachusetts 02115 and * Biogen Research Corp, Cambridge, Massachusetts 02142, USA

The manner in which a membrane protein is anchored to the lipid bilayer may have a profound influence on its function. Most cell surface membrane proteins are anchored by a membrane-spanning segment(s) of the polypeptide chain, but another type of anchor has been described for several proteins: a phosphatidyl inositol glycan moiety, attached to the protein C terminus[1,2]. This type of linkage has been identified on membrane proteins involved in adhesion[3] and transmembrane signalling[4,5] and could be important in the execution of these functions. We report here that an immunologically important adhesion glycoprotein, lymphocyte function-associated antigen 3 (LFA-3), can be anchored to the membrane by both types of mechanism. These two distinct cell-surface forms of LFA-3 are derived from different biosynthetic precursors. The existence of a phosphatidyl-inositol-linked and a transmembrane anchored form of LFA-3 has important implications for adhesion and transmembrane signalling by LFA-3.

LFA-3 is a cell-surface glycoprotein found on erythrocytes, epithelial cells, endothelial cells, fibroblasts and most cells of haematopoietic origin[6]. LFA-3 interacts with the T lymphocyte CD2 membrane glycoprotein, and this ligand-receptor pair mediates intercellular adhesion between LFA-3$^+$ cells and thymocytes, natural killer cells, cytolytic T lymphocutes, and other mature T lymphocytes[6-11]. Cell surface LFA-3 and LFA-3 reconstituted into artificial membranes can both activate T lymphocytes in conjunction with other signals[12,13], which is consistent with the ability of pairwise combinations of anti-CD2 monoclonal antibody (MAb) to activate CD2$^+$ cells[6]. Recently, we have found that LFA-3 is deficient in affected erythrocytes in patients with paroxysmal nocturnal haemoglobinuria (PNH) (ref. 14), an acquired disorder affecting phosphatidyl inositol (PI)-linked proteins[15]. This suggests that LFA-3 is attached to the surface of human erythrocytes by a PI-glycan moeity.

We first obtained evidence for distinct forms of LFA-3 when studying its biosynthesis in the JY B lymphoblastoid cell line. Labelling of JY cells for one minute with [^{35}S]methionine, followed by a five minute chase and isolation with anti-LFA-3-Sepharose, revealed two distinct LFA-3 precursors of relative molecular mass (M_r) of 41,000 (41K) and 37,000 (37K) (Fig. 1*a*, lane 2 arrows). Chase for 10 and 20 minutes showed a gradual decrease in size of the precursors to 39K and 35K (Fig. 1*a*, lanes 3 and 4), which is probably due to trimming of glucose and mannose residues from high-mannose oligosaccharides[16]. There was no apparent interconversion of the two precursors, even during biosynthesis at 24 °C. After chase for 20-40 minutes, the precursors were converted to the mature form of LFA-3, which migrated as a broad band of mean size 65K (Fig. 1*a*, lanes 4 and 5) and corresponded to the form surface-labelled with ^{125}I (ref. 6).

Endoglycosidase H (Endo H) treatment of the 41 and 37K LFA-3 precursors resulted in two bands of corresponding intensity of 29K (p29) and 25.5K (p25.5) respectively (Fig. 1*b*, lanes 1 and 2), but had no effect on mature LFA-3 (Fig. 1*b*, lanes 5 and 6). Therefore each LFA-3 precursor contains 12K of high mannose *N*-linked oligosaccharides which are converted to endo H-resistant complex *N*-linked oligosaccharides during glycoprotein maturation. *N*-glycanase treatment converted both the precursor and mature forms of LFA-3 to two bands at 29K

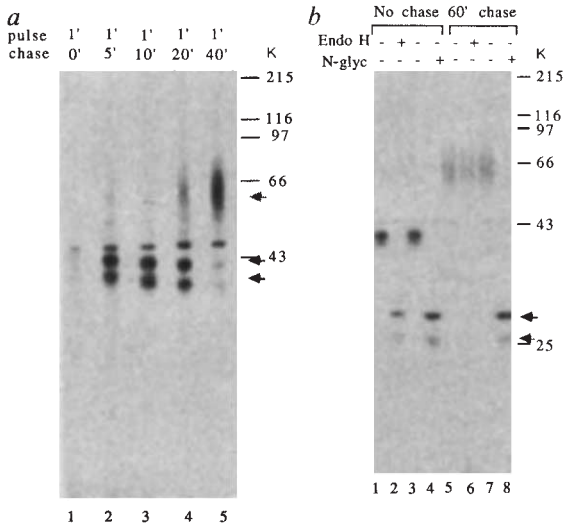


Fig. 1 Biosynthesis of LFA-3. *a*, JY cells were pulsed with ^{35}S-methionine and chased for the indicated times. LFA-3 samples were subjected to reducing SDS-9%-PAGE and fluorography. Solid arrows indicate LFA-3 precursors, hatched arrow indicates mature LFA-3. *b*, As in *a*, except JY cells were pulsed for 5 min and chased as indicated and SDS-denatured LFA-3 samples were incubated with or without 12 mU ml^{-1} Endo H (ref. 26) or 2.5U ml^{-1} *N*-glycanase (*N*-gly)[17] in 50 μl and run on reducing SDS-12%-PAGE. Solid arrow indicates p29, hatched arrow indicates p25.5. The band at 45K on both gels is consistent with actin and was present in MAb-Sepharose isolates of other membrane proteins.

Methods. 5 × 10^6 JY cells were labelled with 0.5 mCi ^{35}S-methionine and lysed[23]. Isolation with anti-LFA-3-Sepharose[11] was as described, except MAb-Sepharose was washed twice with buffer containing 1% Triton X-100, 1% deoxycholate and 0.1% SDS (ref. 24) after incubation with the lysate and before continuing with the washing procedure[25]. SDS-PAGE and fluorography were performed as described previously[23].

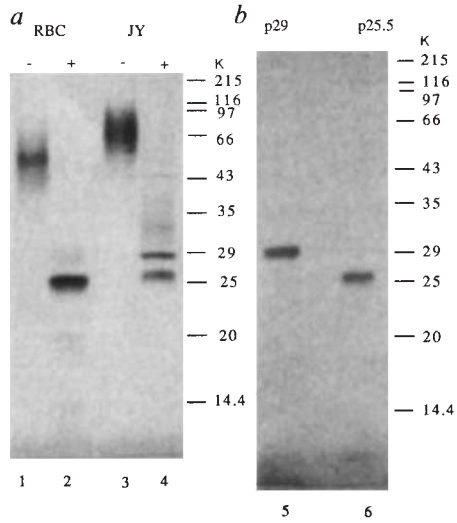


Fig. 2 Purification of LFA-3. LFA-3 was immunoaffinity purified[11] from red blood cells (RBC) (lanes 1 and 2) or JY cells (lanes 3 and 4) and treated with *N*-glycanase[17] (lanes 2 and 4). p29 and p25 from preparative SDS-PAGE of LFA-3, which had been purified from JY cells and treated with *N*-glycanase, were applied to lanes 5 and 6.

Methods. The purification procedure[11] was modified for JY in that 50 g of JY cells were lysed in 500 ml of 2% Triton X-100 in buffer[11] and centrifuged sequentially for 10 min. at 2,000*g* and 2 h at 150,000*g* before affinity chromatography of the supernatant[11]. Preparative PAGE and electroelution were as described[17]. SDS-12%-PAGE with silver staining[11] used 100–500 ng protein.

and 25.5K (Fig. 1*b*, lanes 3,4 and 7,8). These bands were essentially identical in size to those from Endo H treatment of the LFA-3 precursors, allowing for the *N*-acetylglucosamine residue left at *N*-glycosylation sites by Endo H, but not by *N*-glycanase. After the 60 minute chase, both forms of LFA-3 were fully sensitive to proteinase K (not shown). Thus, as shown by studies on their p29 and p25.5 backbones, both the 41K and 37K precursors yield distinct forms of mature, cell surface LFA-3.

The 41K and 37K precursors were seen in similar relative amounts in all cell lines examined, despite the wide size variation in the mature form of LFA-3: Reid-Sternberg line L428, 64K; plasmacytoma line U266B$_1$, 43K; cervical carcinoma line HeLa, 59K; EBV-transformed B lymphoblastoid line F2B, 76K; T lymphoma line Jurkat, 64K; concanavalin A T lymphoblasts, 68K and adherent peripheral blood mononuclear cell fraction, 55K (not shown). The size heterogeneity of mature LFA-3 was due to *N*-linked carbohydrate, as shown with *N*-glycanase; this heterogeneity caused the 2 forms of mature LFA-3 to partially overlap in SDS-PAGE. The size of LFA-3 determined in reducing and non-reducing gels is similar, so LFA-3 does not exist in a disulphide-linked complex on the cell surface (not shown). Immunoaffinity isolation after denaturation in SDS at 100 °C (Fig. 3*b* legend) showed that both p29 and p25.5 possess the TS2/9 (LFA-3 MAb) epitope.

LFA-3 purified[11] in 100 μg quantities from JY B lymphoblastoid cells yielded both p29 and p25.5 after *N*-glycanase digestion (Fig. 2, lanes 3 and 4), as predicted by our biosynthesis studies. In contrast, LFA-3 purified from erythrocytes yielded p25.5 after deglycosylation, and only a trace amount of material comigrating with p29 (Fig. 2, lanes 1 and 2). To determine the relatedness of p29 and p25.5 from JY cells, they were separated by preparative SDS-PAGE (Fig. 2, lane 5 and 6) and 50–80 pmol was subjected to gas phase *N*-terminal sequencing[17]. The *N*-terminal sequence FSQQIYGVVYGDVTFHVPSNVPL was obtained for both p29 and p25.5, and is identical to the *N*-terminal sequence previously described for *N*-glycanase-treated erythrocyte LFA-3 (ref. 17). Therefore, p29 and p25.5 are distinct forms of LFA-3 which share a common epitope and *N*-terminal sequence.

We determined whether one of these forms was selectively released from the surface of JY cells by phosphatidyl inositol-specific phospholipase C (PIPLC). Surface iodinated JY cells were treated with or without PIPLC. LFA-3 remaining on the cell surface was isolated with anti-LFA-3-Sepharose from detergent lysates of the cell pellets and LFA-3 released by PIPLC was isolated from the medium. A portion of cell surface LFA-3 was specifically removed by PIPLC (Fig. 3, lanes 1–4). To resolve which form(s) was released, aliquots of the isolated LFA-3 were digested with *N*-glycanase before SDS-PAGE (Fig. 3, lanes 5–8). PIPLC treatment resulted in no change in the amount of cell surface p29 and released no p29 into the supernatant. In contrast, PIPLC greatly decreased the amount of cell surface p25.5 and a strong band at 26K appeared in the supernatant. The slight change in the mobility of p25.5 after removal of phosphatidylinositol is consistent with reports on other PI-anchored proteins[15].

The PIPLC sensitivity of p25.5 shows that it is a PI-anchored form of LFA-3, but only indirectly indicates that p29 has a polypeptide membrane anchor. The deduced amino-acid sequence of an LFA-3 complementary DNA clone[17] predicts a signal sequence, an 187-amino-acid extracellular domain, a 23-amino-acid transmembrane domain, and a 12-amino-acid cytoplasmic domain. The cytoplasmic tail has potential trypsin cleavage sites 4, 8, and 11 amino acids from the C terminus. In microsomes, the LFA-3 C terminus is predicted to be external. Cells were labelled with [^{35}S]-methionine for 4 minutes, broken in a dounce homogenizer, and the trypsin sensitivity of the LFA-3 precursors in microsomes was assessed[18]. Trypsin treatment of intact microsomes resulted in a 0.5K decrease in p29

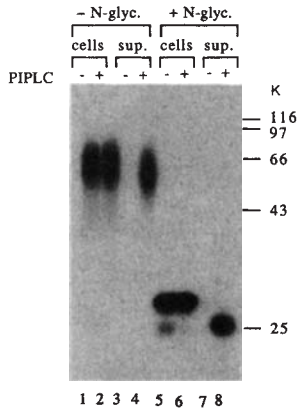


Fig. 3 PI-anchored form of LFA-3. LFA-3 immunoprecipitated from cells treated with or without PIPLC was treated with or without *N*-glycanase, as indicated. 'Cells' indicates LFA-3 from detergent-lysed cells; 'sup' indicates LFA-3 from the media in which the cells were treated with or without PIPLC.LFA-3 immunoaffinity isolated from Triton X-100 lysates of surface iodinated cells (lane 5) had a higher ratio of p29 to p25.5 compared to lysates of cells metabolically labelled for less than 2 h. Greater recovery of LFA-3 after solubilization with PIPLC could be due to more efficient solubilization or immunoaffinity isolation. Resistance of PI-linked proteins in the plasma membrane to solubilization with Triton X-100 has been reported[26] supporting the latter.
Methods. JY cells were iodinated, washed, and treated with *S. aureus* PIPLC (ref. 14). Immunoprecipitation was from lysates of washed cells of from 100,000*g* supernatants of the media, as described. Other methods as in Fig. 1, except that autoradiography was used instead of fluorography[14].

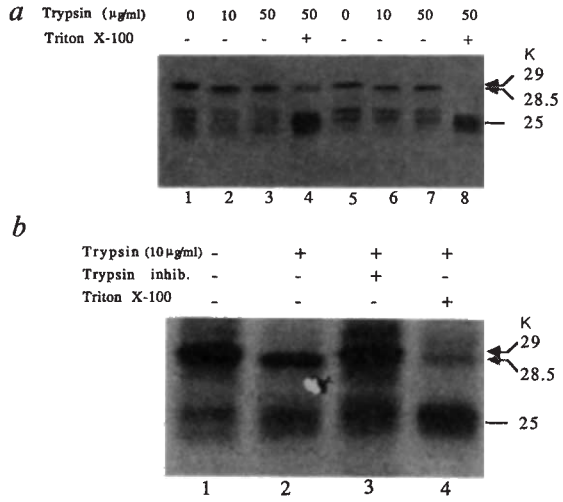


Fig. 4 Transmembrane polypeptide anchored form of LFA-3. *a*, and *b*, LFA-3 precursors, isolated by two cycles (*a*) or one cycle (*b*) of immunoprecipitation from microsomes treated as indicated, were subjected to SDS-12%-PAGE.
Methods. L428 cells (5×10^7) were pulsed for 4 min with 4 mCi of ^{35}S methionine and Dounce-homogenized in a hypotonic buffer[18] containing 2 mM dithiothreitol. The post-nuclear supernatant was treated with or without 10 or 50 μg ml^{-1} trypsin, 0.5 mg ml^{-1} soy bean trypsin inhibitor and/or 0.1% Triton X-100 for 1 h at 4 °C (*a*, lanes 1-4; *b*) or 24 °C (*a*, lanes 5-8) and the reaction was terminated with soy bean trypsin inhibitor. LFA-3 was subjected to a single cycle of isolation with MAb-Sepharose (*b*) or to an additional cycle of isolation after elution from MAb-Sepharose with 0.5% SDS and addition of a sevenfold excess of Triton X-100 (*a*) (ref. 27). The second immunoprecipitation step was to reduce background bands. Digestion with *N*-glycanase and SDS-12%-PAGE and fluorography as in Fig. 1.

(Fig. 4*a* and *b*, lane 2, compare lane 1). The apparent size decrease in p29 was blocked by prior addition of soy bean trypsin inhibitor to the post-nuclear supernatant (Fig. 4*b*, lane 3). When detergent was included, the bands at 28.5K and 25.5K were decreased in intensity or eliminated (Fig. 4*a*, lanes 4 and 8, Fig. *b*, lane 4) and a band of 24–25K appeared instead. There was no shift detected in p25.5 in the absence of Triton X-100. The trypsin sensitivity of the p29 form of LFA-3 in intact microsomes strongly suggests that it is a transmembrane protein.

We have found two cell-surface forms of LFA-3 which have distinct PI and polypeptide chain membrane anchors and are derived from distinct biosynthetic precursors. Our results are concordant with findings on LFA-3 in PNH. PNH, an acquired clonal defect in hematopoietic cells, leads to a selective absence of PI-anchored proteins[1,15]. Our result that p25.5 has a PI membrane anchor is in agreement with the finding that the erythrocyte form of LFA-3 corresponds to p25.5, and that LFA-3 is deficient on PNH erythrocytes[14]. In contrast, neutrophils and monocytes from PNH patients which are clonally affected, as shown by complete deficiency of decay accelerating factor, express only slightly reduced quantities of LFA-3 (M. E. Medof, P.S., M.L.D., D. G. Ayers, E. I. Walter, H. A. Stafford, R. Green, M. L. Tykocinski and T.A.S., unpublished). We predict that the form of LFA-3 on the surface of affected leukocytes from PNH patients is p29. p29 may correspond to LFA-3 exactly as encoded by a recently isolated LFA-3 cDNA (ref. 17), whereas p25.5 might be encoded by the same messenger RNA or a different mRNA, possibly generated by alternative splicing, as reported for neural cell adhesion molecule[19].

In addition to mediating intercellular adhesion by binding to CD2, LFA-3 might be able to transduce signals, as the binding of anti-LFA-3 MAb to monocytes, thymic epithelial cells and keratinocytes enhances the release of IL-1 (ref. 20). The presence of PI- and polypeptide-anchored forms of LFA-3 on the surface of nucleated cells raises the possibility of division of the adhesion and signalling functions between these two form of LFA-3. The PI-anchored form is predicted to have a significantly faster lateral mobility in the cell membrane[21], which would facilitate diffusion of LFA-3 to sites of adhesion with CD2$^+$ cells[22]. Expression of PI- and polypeptide anchored form of neural cell adhesion molecule has also been reported[3,19], suggesting that the expression of ligand binding domains of adhesion molecules with different types of membrane anchorage may be a general strategy for regulating adhesiveness or other functions.

We thank Dr Martin Low for his gift of *S. aureus* PIPLC, and Ms Cynthia A. Burne for *N*-terminal sequence analysis. This work was supported by an NIH grant.

Received 30 July; accepted 15 September 1987.

1. Low, M. G. *Biochem. J.* **244**, 1-13 (1987).
2. Cross, G. A. M. *Cell* **48**, 179-181 (1987).
3. He, H. T., Barbet, J., Chaix, J. C. & Goridis, C. *EMBO J.* **5**, 2489-2494 (1986).
4. Kroczek, R. A., Gunter, K. C., Germain, R. N. & Shevach, E. M. *Nature* **322**, 181-184 (1986).
5. Reiser, H. *et al. Cell* **47**, 365-370 (1986).
6. Springer, T. A., Dustin, M. L., Kishimoto, T. K. & Marlin, S. D. *A. Rev. Immun.* **5**, 223-252 (1987).
7. Vollger, L. W., Tuck, D. T., Springer, T. A., Haynes, B. F. & Singer, K. H. *J. Immun.* **138**, 358-363 (1987).
8. Plunkett, M. L., Sanders, M. E., Selvaraj, P., Dustin, M. L. & Springer, T. A. *J. exp. Med.* **165**, 664-676 (1987).
9. Makgoba, M. W., Shaw, S., Gugel, E. A. & Sanders, M. E. *J. Immun.* **138**, 3587-3589 (1987).
10. Selvaraj, P. *et al. Nature* **326**, 400-403 (1987).
11. Dustin, M. L., Sanders, M. E., Shaw, S. & Springer, T. A. *J. exp. Med.* **165**, 677-692 (1987).
12. Hunig, T., Tiefenthaler, G., Meyer zum Buschenfelde, K.-H. & Meuer, S. C. *Nature* **326**, 298-301 (1987).
13. Bierer, B. E. *et al. Fedn Proc.* **46**, 1498 (abstr. 6894) (1987).
14. Selvaraj, P., Dustin, M. L., Silber, R., Low, M. G. & Springer, T. A. *J. exp. Med.* **166**, 1011-1025 (1987).
15. Davitz, M. A., Low, M. G. & Nussenzweig, V. *J. exp. Med.* **163**, 1150-1161 (1986).
16. Kornfeld, R. & Kornfeld, S. *A. Rev. Biochem.* **54**, 631-664 (1985).
17. Wallner, B. P. *et al. J. exp. Med.* **166**, 923-932 (1987).
18. Braell, W. A. & Lodish, H. F. *J. biol. Chem.* **256**, 11337-11344 (1981).
19. Barthels, D. *et al. EMBO J.* **6**, 907-914 (1987).
20. Le, P., Denning, S., Springer, T., Haynes, B. & Singer, K. *Fedn Proc.* **46**, 447 (1987).
21. Ishihara, A., Hou, Y. & Jacobson, K. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1290-1293 (1987).
22. Singer, S. J. In *Surface membrane receptors. Interface between cells and their environment*, (eds Bradshaw, R. A., Frazier, W. A., Merrell, R. C. & Gottlieb, D. I.) 1-24 (Plenum, New York 1976).
23. Sastre, L., Kishimoto, T. K., Gee, C., Roberts, T. & Springer, T. A. *J. Immun.* **137**, 1060-1065 (1986).
24. Gielkens, A. L. J., Van Zaane, D., Bloemers, H. P. J. & Bloemendal, H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 356-360 (1976).
25. Kürzinger, K. & Springer, T. A. *J. biol. Chem.* **257**, 12412-12418 (1982).
26. Hoessli, D. & Rungger-Brandle, E. *Expl cell Res.* **156**, 239-250 (1985).
27. Lublin, D. M., Krsek-Staples, J., Pangburn, M. K. & Atkinson, J. P. *J. Immun.*. **137**, 1629-1635 (1986).

Growth inhibition by protein kinase C late in mitogenesis

Chou-Long Huang & Harlan E. Ives

Department of Physiology and Cardiovascular Research Institute, University of California, San Francisco, California 94143, USA

The importance of α-thrombin in the clotting cascade is well-known, but it is also a potent mitogen[1–3]. Like many other mitogens, thrombin causes receptor-mediated activation of a phosphatidylinositol-specific phospholipase C (PLC), leading to the release of diacylglycerol and the subsequent activation of protein kinase C (refs 3–6). Protein kinase C is probably important in cell proliferation, as activation of this enzyme by phorbol esters promotes growth in many systems[4–6]. Some growth factors have tyrosine kinase activity and function without activation of PLC or protein kinase C[7–9]. In this report we show that α-thrombin retains its mitogenicity in vascular smooth muscle cells depleted of protein kinase C. Phorbol-12-myristate-13-acetate (PMA) is found to be a potent growth inhibitor when added to vascular smooth muscle cells with α-thrombin. Moreover, growth inhibition is maximal when protein kinase C is activated 4 hours after exposure to thrombin, long after the completion of 'early events' induced by thrombin. Thus, PMA probes an event late in the G_1 phase of the cell cyle or at the G_1–S transition.

Activation of protein kinase C by phorbol esters or 1,2-diacylglycerols exerts both positive and negative effects on cellular events associated with mitogenesis[6,10]. Protein kinase C can stimulate Na^+/H^+ exchange, an activity associated with mitogenicity in many cell types[10], including vascular smooth muscle cells[11]. On the other hand, we found that exposure for 10 min to PMA at 200 ng ml^{-1} caused 82% inhibition of thrombin-induced inositol trisphosphate release in vascular smooth muscle cells, from 1,110 ± 78 c.p.m. per dish to 156 ± 15 c.p.m. per dish ($n = 4$) when measured (for method, see ref. 12) 30 s after addition of thrombin (0.4 U ml^{-1}). PMA also blunted the thrombin-induced Ca^{2+} spike (Fig. 1). These effects of PMA must require protein kinase C activity as they were absent in protein kinase C-depleted cells (Fig. 1). To see if PMA was blocking these early cellular events through interference with thrombin binding, we measured ^{125}I-labelled thrombin binding as previously described[13]. Exposure for 30 minutes to PMA did not alter either the binding affinity (5.8 nM in control; 6.2 nM with PMA) or the maximum binding of thrombin (210 ng per dish in the control, compared with 240 ng per dish with PMA). Lastly, there was no effect of PMA pretreatment on platelet-derived growth factor (PDGF)-induced inositol phosphate release or Ca^{2+} transients (data not shown), indicating that the blunting of these events is specific for thrombin, or a class of agents separate from PDGF.

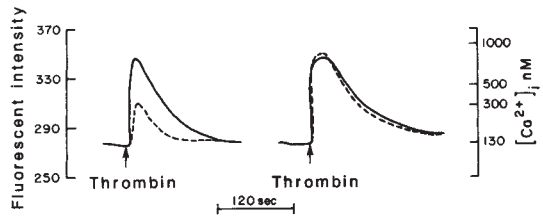


Fig. 1 Activation of protein kinase C blunts the thrombin-induced Ca^{2+} spike. Control cells (left) were treated with 200 ng ml^{-1} PMA (dotted line) or without PMA (solid line) 10 min before addition of 0.4 U ml^{-1} thrombin. Protein-kinase-C-depleted cells (right) were rechallenged with 200 ng ml^{-1} PMA (dotted line) 10 min before addition of 0.4 U ml^{-1} thrombin; solid line, cells without PMA. Fluorescent intensity is in arbitrary units and the excitation wavelength was 340 nm.

Methods. Neonatal rat vascular smooth muscle cells (strain R22) were obtained from Peter Jones[24]. Cells were maintained in Eagle's medium with 10% fetal calf serum, 2% tryptose phosphate broth, penicillin (50 U ml^{-1}), and streptomycin (50 U ml^{-1}) in a humidified atmosphere of 5% CO_2 95% air at 37 °C and were used for the present studies at passages 11–17. Cells were grown to confluence and intracellular Ca^{2+} was measured using fura 2 (ref. 25). Protein kinase C was depleted by 24 h incubation with 250 ng ml^{-1} of PMA as previously described in fibroblasts[17]. In the vascular smooth muscle cells used for this study, PMA (250 ng ml^{-1} for 24–72 h) similarly depletes functional protein kinase C activity (assayed by phorbol ester activation of Na^+/H^+ exchange, ref. 11). Using this assay, we found that treatment with 3 ng ml^{-1} PMA (12 h) also causes total loss of functional protein kinase C activity.

Table 1 DNA synthesis induced by thrombin but not by PDGF is inhibited by PMA

	^{3}H-thymidine incorporation (c.p.m. per dish)	
Control cells		
Basal	2,903 ± 162	
PMA	2,832 ± 150	
Thrombin	25,555 ± 1,480	$P < 0.01$
PMA + thrombin	11,663 ± 723	
PDGF	12,016 ± 973	NS
PMA + PDGF	12,305 ± 526	
Protein-kinase-C-depleted cells		
Basal	2,982 ± 218	
PMA	2,599 ± 258	
Thrombin	19,359 ± 994	NS
PMA + thrombin	18,117 ± 816	

Confluent cultures of vascular smooth muscle cells were grown in 24-well plates and growth-arrested by placing them in medium containing 20 mM HEPES (pH 7.4), 5 μg ml^{-1} transferrin, and 0.5 mg ml^{-1} bovine serum albumin for 72 h. Cells were depleted of protein kinase C by simultaneous 72-h incubation with 250 ng ml^{-1} PMA (Fig. 1, legend). Cells were exposed to thrombin (1 U ml^{-1}), PDGF (7.5 nM), or PMA (100 ng ml^{-1}) and 1 μCi of ^{3}H-thymidine per well for 24 h. If PMA was used, it was added 10 min before thrombin or PDGF. Cells were fixed with trichloroacetic acid (TCA) and TCA-insoluble counts were determined. Ethanol used for PMA solution had no effect on thrombin-induced DNA synthesis.Each data point is the mean ±s.e.m. of quadruplicate determinations in one of six experiments. NS, not significant.

Although the net effect of PMA on cell growth is positive in most systems[6,10], there are examples of growth inhibition by phorbol esters[14,15]. In vascular smooth muscle cells, we found that PMA inhibited DNA synthesis induced by thrombin, but not that induced by PDGF. (Table 1 top), and that PMA alone had no effect on DNA synthesis. Like the effect of PMA on the early events associated with thrombin activation, protein kinase C depletion obliterated the growth inhibitory effect of acute PMA administration (Table 1 bottom). Thus, PMA appears to interfere specifically with both early biochemical events and DNA synthesis induced by thrombin.

To see if the blunting of early events could explain the growth inhibition observed with PMA in vascular smooth muscle cells, we added PMA to cultures at different times before and after addition of thrombin (Fig. 2). Surprisingly, PMA did not cause maximal inhibition of thrombin-induced DNA synthesis when it was added before thrombin; rather, inhibition was greatest when PMA was added after thrombin, reaching 100% when it was added 4 h after thrombin. Later addition of PMA resulted in restoration of thrombin-induced DNA synthesis. As in Table 1, PMA did not inhibit PDGF-induced ^{3}H-thymidine incorporation, even when added 4 h after PDGF (see Fig. 3 legend). The dose-response curve for inhibition by PMA of thrombin-

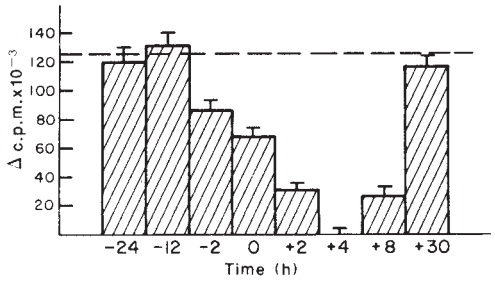


Fig. 2 Maximal inhibition of thrombin-induced DNA synthesis by PMA 4 h after exposure to thrombin. Cells were prepared as described in Table 1, and incubated with thrombin and ^{3}H-thymidine simultaneously for 36 h. PMA (100 ng ml^{-1}) was added to the wells at the indicated time relative to the addition of thrombin so −2 h represents addition of PMA 2 h before thrombin and +2 h represents addition 2 h after thrombin. Incorporation of ^{3}H into TCA-insoluble form was determined; data is presented as the difference in thymidine incorporation with thrombin and PMA above basal incorporation. Dotted line, ^{3}H incorporation in cells incubated with thrombin alone for 36 h.

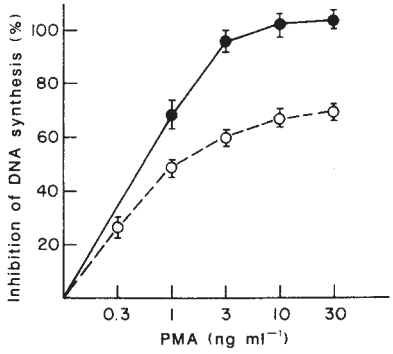


Fig. 3 Dose-response curve for inhibition by PMA of thrombin-induced DNA synthesis. ^{3}H-thymidine incorporation into cells incubated with thrombin (1 U ml^{-1}) for 36 h was determined as in Fig. 2. PMA at the indicated dose was added either simultaneously (dotted line) or 4 h after (solid line) addition of thrombin. Data is presented as % inhibition of thrombin-induced thymidine incorporation by PMA. When added 4 h after mitogen, PMA (3 ng ml^{-1}) inhibited thrombin-induced thymidine incorporation by 96% but failed to inhibit PDGF-induced thymidine incorporation at a maximal dose of PDGF (7.5 nM) as used in Table 1 (7,979 ± 177 c.p.m. per dish in control; 28,708 ± 1,116 for PDGF; 33,722 ± 1,816 for PDGF with PMA added 4 h later).

induced ^{3}H-thymidine incorporation is shown in Fig. 3. The dose of PMA required for half-maximal inhibition of ^{3}H-thymidine incorporation was ~1 ng ml^{-1}, which agrees closely with the dose required for half-maximal activation of protein-kinase-C-mediated Na^+/H^+ exchange (0.01 *p*H unit min^{-1} for 0.3 ng ml^{-1} PMA; 0.02 *p*H unit min^{-1} for 3 ng ml^{-1} PMA; 0.02 *p*H unit min^{-1} for 200 ng ml^{-1} PMA). Comparable dose-response curves for activation of protein kinase C by PMA have been observed in platelets[16] and fibroblasts[17].

PMA is not rapidly degraded by biological systems[6] so at first it was not clear why its effect on thrombin-induced mitogenesis was maximal when added 4 h after thrombin. The best explanation is that the growth inhibition by PMA requires fully active protein kinase C and that addition of PMA before thrombin resulted in significant depletion of protein kinase C activity 2-6 h later. Consistent with this idea, PMA caused total loss of functional protein kinase C within 12 h in our system (data not shown). Furthermore, the half-time for protein kinase C depletion in cells exposed to PMA is 6 h in fibroblasts[17] and 30 min in keratinocytes[18]. Thus, maximum growth inhibition in vascular smooth muscle cells results when PMA is added close to the critical time point for its action.

Although recent work has mainly focused on growth stimulation, growth inhibition is also important in the development of normal tissues, in wound-healing, and in tumour formation[19]. Several growth inhibitors which could mediate these processes have been identified, including transforming growth factor-β (ref. 20) and heparin[21], but their mechanism of action is poorly understood. Our results identify the activation of protein kinase C late in G_1 as a potential target for growth inhibitors.

These and other results[6,10] indicate that protein kinase C could have several functions in the regulation of cell growth. Usually PMA is synergistic for growth when used simultaneously with another mitogen[6,10], but here PMA inhibits growth when it is applied 4 h later than mitogen. Although PMA does not stimulate growth at any time in our system, activation of protein kinase C could stimulate and inhibit mitogenesis in some cells. The net effect of stimulating protein kinase C would depend on its timing or on the intensity of these opposing processes.

We believe this to be the first reported action of protein kinase C removed in time from the initial binding of a mitogen. Although the function of protein kinase C at this distal time point is not known, our data suggest that it might interfere with continuing action of the thrombin receptor. Thrombin must remain bound to its receptor for 8 h to achieve maximal synthesis of DNA[22], whereas PDGF needs only 30 min (ref. 23). Our results show that PMA interferes with thrombin-, but not with PDGF-induced events both immediately (Fig. 1) and 4 h after (Fig. 2) addition of hormone. Thus, the cellular response to mitogenic stimulation seems to be specific for the original mitogen for at least 4 h.

We thank Floyd C. Rector Jr and R. Curtis Morris for support and Henry R. Bourne and Lewis T. Williams for critical reading of the manuscript. We also thank John Fenton III and Lewis T. Williams for supplying us with α-thrombin and PDGF respectively. This work was supported in part by the NIH.

Received 11 June; accepted 14 September 1987.

1. Carney, D. H., Stiernberg, J. & Fenton, J. W. *J. cell. Biochem.* **26**, 181-195 (1984).
2. Carney, D., Scott, D., Gordon, E. & LaBelle, E. *Cell* **42**, 479-488 (1985).
3. Chambard, J. C., Paris, S., L'Allemain, G. & Pouyssegur, J. *Nature* **326**, 800-803 (1987).
4. Dicker, P. & Rozengurt, E. *Nature* **287**, 607-612 (1980).
5. Berenblum, I. in *Cancer, A Comprehensive Treatise* Vol. 1 (ed. Becker, F. F.) 323-344 (Plenum, New York, 1975).
6. Nishizuka, Y. *Science* **233**, 305-312 (1986).
7. Besterman, J., Watson, S. & Cuatrecasas, P. *J. biol. Chem.* **261**, 723-727 (1986).
8. Coughlin, S. R., Lee, W. F., Williams, P. W., Giels, G. M. & Williams, L. T. *Cell* **43**, 243-251 (1985).
9. Magnaldo, I. *et al. J. biol. Chem.* **261**, 16916-16922 (1986).
10. Rozengurt, E. *Science* **234**, 161-166 (1986).
11. Huang, C.-L., Cogan, M. G., Cragoe, E. J. & Ives, H. E. *J. biol. Chem.* **262**, 14134-14140 (1987).
12. Berridge, M. J., Downes, P. C. & Hanley, M. R. *Biochem. J.* **206**, 587-595 (1982).
13. Carney, D. H. & Cunningham, D. D. *Cell* **15**, 1341-1346 (1978).
14. Doctrow, S. R. & Folkman, J. *J. Cell Biol.* **104**, 679-687 (1987).
15. Huberman, E. & Callahan, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1293-1297 (1979).
16. Castagna, M., Takai, Y., Kaibuchi, K., Kikkawa, U. & Nishizuka, Y. *J. biol. Chem.* **257**, 7847-7851 (1982).
17. Rodriguez-Pena, A. & Rozengurt, E. *Biochem. biophys. Res. Commun.* **120**, 1053-1059 (1984).
18. Chida, K., Kato, N. & Kuroki, T. *J. biol. Chem.* **261**, 13013-13018 (1986).
19. Marx, J. L. *Science* **232**, 1093-1095 (1986).
20. Tucker, R. F., Shipley, G. D., Moses, H. L. & Holley, R. W. *Science* **226**, 705-707 (1984).
21. Clowes, A. W. & Karnovsky, M. J. *Nature* **265**, 625-626 (1977).
22. Van Obberghen-Schilling, E., Perez-Rodriguez, R. & Pouyssegur, J. *Biochem. biophys. Res. Commun.* **106**, 79-86 (1982).
23. Pledger, W. J., Stiles, C. D., Antoniades, H. N. & Scher, C. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4481-4485 (1987).
24. Jones, P. A., Scott-Burden, T. & Gevers, W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 353-357 (1979).
25. Ives, H. E. & Daniel, T. O. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1950-1954 (1987).

Novel chimaeric protein expressed in Philadelphia positive acute lymphoblastic leukaemia

Lesley C. Walker*, Trivadi S. Ganesan*, Susheela Dhut*, Barbara Gibbons*, T. Andrew Lister*, Jonathan Rothbard† & Bryan D. Young*‡

* Imperial Cancer Research Fund, Department of Medical Oncology, St Bartholomew's Hospital, West Smithfield, London EC1, UK
† Laboratory of Molecular Immunology, Imperial Cancer Research Fund, Lincoln's Inn Fields, London, UK

Cytogenic changes are becoming increasingly important in understanding the pathogenesis of human malignancies. The t(9;22) (q34;q11) translocation is one of the most consistent and generates the Philadelphia chromosome (Ph[1]) (ref. 1) in chronic myeloid leukaemia (CML); it has also been observed in some acute lymphoblastic leukaemias (ALL) (ref. 2). In CML the breakpoints occur on chromosome 22 in the region designated *bcr* (ref. 3) and result in the expression of a bcr–abl fusion product of relative molecular mass (M_T) 210,000 (210K) with associated *in vitro* tyrosine kinase activity[4] (P210[bcr–abl]). In some cases of Ph[1]-positive ALL, a novel abl-related protein (P190[all–abl]) of 190K has been shown[5–7] to have tyrosine kinase activity. In this report we demonstrate that the P190[all–abl] protein has a bcr determinant from the amino-terminal region, but is lacking a bcr determinant normally found in the P210[bcr–abl] near the bcr-abl junction. The chimaeric nature of the P190[all–abl] was confirmed by sequential immunoprecipitation with antisera against abl and bcr peptides.

About 20% of cases of adult ALL and about 5% of childhood ALL carry a t(9:22) translocation[8] that is cytogenetically indistinguishable from the Philadelphia translocation characteristic of CML. It has therefore been difficult to differentiate a *de novo* Ph[1]-positive ALL from a CML lymphoid blast crisis, which may have had an undetected asymptomatic chronic phase. It was found[9] that in two Ph[1]-positive ALL patients the breakpoints on chromosome 22 had occurred outside the bcr region which is normally broken in CML (ref. 3) suggesting a molecular basis for such a distinction. In further studies it was shown that there is a subgroup of Ph[1]-positive ALL patients with no observable rearrangements at the *bcr* locus[10], and that the breakpoints on chromosome 22 lie between this region and the more proximal *Cλ* locus[11]. There seems to be a correlation between the absence of a detectable rearrangement in the CML-specific breakpoint region and the presence of the P190[all-abl] in Ph[1] positive ALL (refs. 5, 6, 7). Although a rearrangement has been reported[12] in a Ph[1]-positive ALL to lie between 20–100 kilobases(kb) 5′ of the *bcr* region, it is still unclear whether the amino terminus of the P190[all-abl] contains any sequence encoded by the *phl* gene, as the complete gene is now designated[13].

‡ To whom correspondence should be addressed.

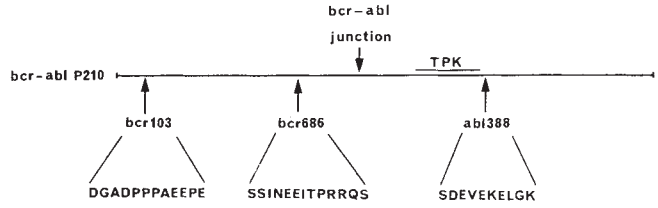


Fig. 1 Sequence and positions on the P210 of the peptides to which anti-sera were raised.
Methods. The bcr peptides, amino acids 103–114 (bcr103) and 686–698 (bcr686), were chosen from the putative translation of bcr cDNA sequence[14] and the abl peptide, amino acids 388–397 (abl388) corresponded to a previously synthesized v-abl sequence[15]. Peptides were synthesized on an Applied Biosystems 430A peptide synthesiser using t-boc chemistry and commercially available reagents. The peptides were cleaved from the resin using HF, partially purified, and analysed by HPLC[24]. For immunization, cysteine was added to the amino terminus of each peptide, enabling it to be conjugated to thyroglobulin and bovine serum albumin[25].

To address this question we prepared rabbit antisera against synthetic peptides from three regions, summarized in Fig 1*a*. The two bcr peptides were from the predicted 5′ bcr amino-acid sequence[14], and the abl peptide was a v-abl sequence previously shown to be antigenic[15]. Cell extracts from the Ph[1] positive cell line BV173, derived from the cells of a patient with CML in blast transformation[16], were immunoprecipitated with each antiserum and the immune complexes labelled in an *in vitro* tyrosine kinase reaction. Each antiserum detected the P210[bcr-abl], whereas antisera inhibited with cognate peptide and control pre-immune sera were negative (data not shown). Identical results were obtained from the CML-derived cell line K562, whereas no P210-labelled species were detected from the myeloid cell line KG1, which lacks the Philadelphia translocation (data not shown). These experiments confirmed the specificity of the anti-peptide sera for the P210[bcr-abl], and the predicted translation of the 5′ region of the *bcr* complementary DNA sequence[17].

DNA from cells of three patients with Ph[1]-positive ALL was analysed by Southern blotting and hybridization with two bcr probes. In the sample from patient 1 a rearrangement was detected by both bcr probes, whereas the samples from patients 2 and 3 showed no rearrangements detectable with either probe and using a range of restriction enzymes (Table 1). Cells from the ALL patients 2 and 3 were therefore expected to express the P190[all-abl] species[5,6,7]. Cell lysates from each patient were immunoprecipitated with the anti-peptide sera and the results of the *in vitro* kinase reaction are shown in Fig. 2*a*. As expected, a P210 was detected from cells of patient 1 with the anti-abl388 serum and with both anti-bcr sera, indicating that this protein was similar to that detected in BV173. The anti-abl388 serum detected a P190 peptide in patients 2 and 3. The anti-bcr103 serum yielded a similar P190 species whereas the anti-bcr686 gave only a faint diffuse signal at this position in the case of patients 2 and 3. These data indicated that the P190[all-abl] in

Table 1 Summary of findings for cell line BV 173 and cells of 3 patients with Ph[1] positive acute lymphoblastic leukaemia

Cells	Phenotype	Karyotype	Bcr* rearr.	Anti-bcr103	Anti-bcr686	Anti-abl388
BV173	Common ALL[28]	46, X, −10, 1q+, i(8q), t(9;22) (q34;q11), inv(12)(p12q12),+Ph′,+mar	+	210K	210K	210K
ALL 1	Common ALL	46, XX, t(9;22) (q34;q11)	+	210K	210K	210K
ALL 2	B Lymphoid	46, XX, t(9;22) (q34;q11)	−	190K	—	190K
ALL 3	Common ALL	55, XXX, +2, +4, +8, +10, +14, +21, +21 t(9;22) (q34;q11), +Ph′	−	190K	—	190K

* An intron *bcr* probe corresponding to the 3′ *Hin*dIII-*Bgl*II 1.2-kb fragment (Oncogene Sciences Inc.) and a *bcr* cDNA probe (0.45 kb *Eco*RI-*Pst*I fragment subcloned from the *bcr-abl* mRNA expressed in K562 cells) were used in Southern analysis with a range of restriction enzymes.

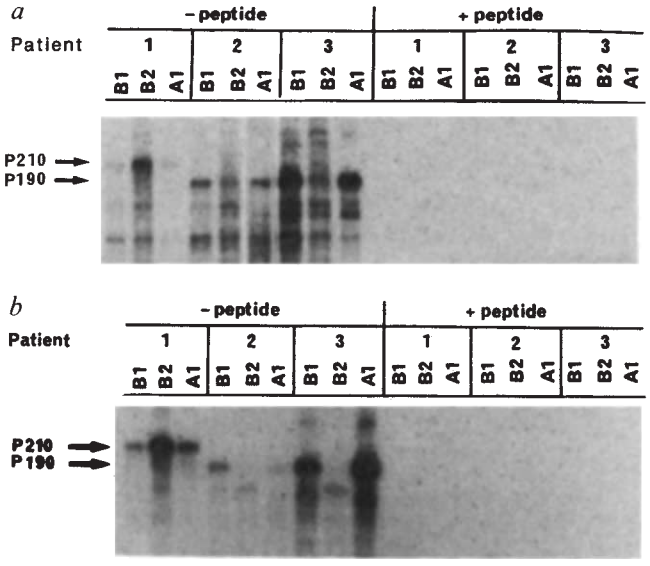


Fig. 2 *In vitro* tyrosine kinase assay on immune complexes from cells of three patients with acute lymphoblastic leukaemia. *a*, The anti-sera are designated as follows: B1, anti-bcr103; B2, anti-bcr686; Al, anti-abl388. Appropriate cognate peptide was added as indicated. *b*, The same gel after alkaline hydrolysis.

Methods. Rabbits were immunized with 2 ml complete Freund's adjuvant containing 100 μg peptide-thyroglobulin conjugate, administered at multiple subcutaneous sites. One or two booster injections of 100 μg antigen in incomplete Freund's adjuvant were given monthly. Antisera were selected for specificity and titre for cognate peptide by solid phase radioimmunoassay against BSA-peptide conjugates, and antibodies purified by affinity chromatography on columns of bovine serum albumin-peptide linked to Sepharose 4B (Pharmacia). Pre-immune IgG was purified on protein A-Sepharose 4B (Pharmacia). For the *in vitro* tyrosine kinase assay, samples were prepared as described[5] and 5×10^6 ml^{-1} cells extracted with lysis buffer (1% Triton X-100, 0.05% SDS, 10 mM sodium phosphate, pH 7.0, 150 mM NaCl, 250 μM phenylmethylsulphonyl fluoride, 70 μg ml^{-1} aprotinin Sigma, 100 μM sodium vandate) for 30 min at 4 °C. Lysates were centrifuged at 10,000*g* for 2 min at 4 °C. Immunoprecipitation with 10 μg affinity-purified antibody was performed as described[5], except that antibody and protein A-Sepharose beads were added simultaneously and the samples rotated gently for 1 h at 4 °C. *In vitro* phosphorylation of washed immunoprecipitates using 10 μCi of (>5,000 Ci $mmol^{-1}$) [γ-^{32}P]ATP (Amersham) and SDS-PAGE were as described[7] with minor modifications.

patients 2 and 3 lacked the bcr686 sequence and contained the bcr103 sequence. Alkaline treatment of the gel selectively hydrolyses phosphoserine and phosphothreonine, but the P190 species detected in the cells from patients 2 and 3 by anti-abl388 and anti-bcr103 remained unaffected, indicating that phosphorylation of this molecule was predominantly on tyrosine (Fig. 2*b*), as reported previously[5]. The 210K species detected from patient 1 with anti-bcr686 serum resisted treatment with alkali, whereas the diffuse bands seen with this antiserum in the case of patients 2 and 3 were selectively removed, suggesting that they were not due to tyrosine phosphorylation. No signal was obtained with sera inhibited with cognate peptides. Together these results indicate that $P190^{all-abl}$ lacks the bcr686 sequence but contains the bcr103 sequence.

Sequential immunoprecipitations were used to establish definitively the chimaeric nature of $P190^{all-abl}$. Both the $P210^{bcr-abl}$ and $P190^{all-abl}$ were first immunoprecipitated with the anti-abl388 serum and then labelled during the *in vitro* kinase reaction. Both labelled species were reprecipitated with each anti-peptide serum and with control sera; the results are shown in Fig. 3. It is clear that the $P210^{bcr-abl}$ was reprecipitated with both anti-bcr sera, confirming that the $P210^{bcr-abl}$ contains the

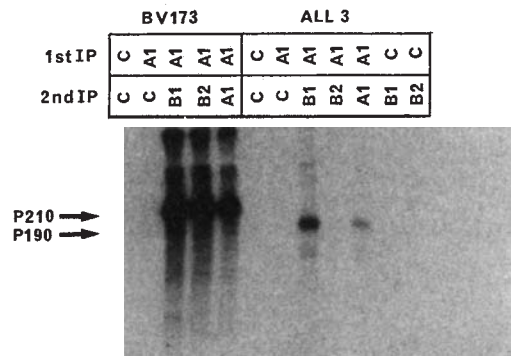


Fig. 3 Sequential immunoprecipitation (IP) of cell lysates from BV173 and ALL patient 3. Each serum is indicated as follows: B1, anti-bcr103; B2, anti-bcr686; A1, anti-abl388; C, control pre-immune serum.

Methods. Treatment of frozen samples and methodology for the immune complex kinase assay was as described[5]. Following the 1st immunoprecipitation step and reaction with [γ-^{32}P]ATP, labelled phosphoproteins were recovered from washed beads by boiling for 3 min in 30 μl 2×SDS sample buffer in the absence of reducing agent and diluted 33-fold in lysis buffer. Samples were reprecipitated for 1 h at 4 °C using 10 μg of anti-peptide antibody or normal rabbit IgG and 20 μl of packed protein A-Sepharose 4B beads. Antigen-antibody complexes were recovered from washed beads and analysed on SDS-polyacrylamide gels as in Fig. 2.

bcr686 and bcr103 sequences. In contrast, the $P190^{all-abl}$ could be reprecipitated only with the anti-bcr103 serum, and not with the anti-bcr686 serum. Reprecipitation of both P210 and P190 with control antisera was negative, proving that the anti-abl388 antibody was completely cleared after the first precipitation.

It is clear that the $P190^{all-abl}$, like the $P210^{bcr-abl}$, is a chimaeric product containing elements of the c-*abl* and *phl* (or *bcr*) genes. Although we cannot exclude a breakpoint in another *bcr*-related gene encoding a sequence identical to bcr103, the simplest explanation is that the breakpoint is in the *phl* gene between the regions encoding bcr103 and bcr686, thus accounting for the lack of rearrangement using conventional bcr probes, and for the absence of bcr686 from $P190^{all-abl}$ (Table 1). The distance between bcr686 and the nearest bcr-abl junction in the $P210^{bcr-abl}$ is 204 amino acids, which is roughly equivalent to the difference between 210K and 190K. Hence if both species contained the same abl component, the bcr-abl junction in $P190^{all-abl}$ may not be far from bcr686 in the amino-terminal direction. The $P190^{all-abl}$ has only been observed in cells of patients classified as Ph^1 positive ALL, and has not been found in well-documented lymphoid blast crisis of CML[5-7]. Although some apparent *de novo* Ph^1 positive ALL clearly express a $P210^{bcr-abl}$ (patient 1; ref. 6), these patients could be typical of a subgroup in which a transition from ALL to CML has been documented[18]. Note that the ALL of patient 3 had some cytogenetic features of blast crisis of CML (an extra Ph^1 chromosome and +8 chromosome) among other changes. The marrow of this patient in clinical remission showed a simple Ph^1 translocation in 20% of metaphases, indicating that the other changes were of a secondary nature. The study of marrow in haematological remission and of the clinical course in ALL patients expressing either species should help to establish whether the $P190^{all-abl}$ is a characteristic of *de novo* Ph^1 positive ALL.

The differences between the $P210^{bcr-abl}$ and $P190^{all-abl}$ and their apparent disease specificities suggest that the nature of their amino terminal alterations are critical to their function. Although it is established[19] that *bcr* exon 3-encoded sequence (24 amino acids) is not necessarily present in the $P210^{bcr-abl}$, there is no apparent clinical correlation with its absence or presence[20]. In contrast, the $P190^{all-abl}$ is lacking a much larger region of bcr (at least 200 amino acids) compared with the

P210$^{bcr-abl}$, and this may be critical for its specificity for *in vivo* substrate or for its subcellular location. There is an important parallel[21] to be drawn from studies of the transforming potential of the P160$^{gag-abl}$ protein of Abelson murine leukaemia virus: the *gag* sequence corresponding to the amino terminus is dispensable for fibroblast transformation, but is needed for stabilizing the v-abl protein in lymphoid cells[22]. Furthermore, myristylation of the amino terminus of P160$^{gag-abl}$ could mediate its location on the inner surface of the cell membrane[23]. This modification has not been detected[21] for the P210$^{bcr-abl}$, and its subcellular location is uncertain. Any differences in the subcellular location or substrate specificity of P210$^{bcr-abl}$ and P190$^{all-abl}$ will be important in interpreting the function of the *bcr*-encoded moieties of these molecules.

We thank Drs E. Canaani for the *bcr* cDNA probe, J. Goldman for the BV173 cell line, M. Horton for phenotype analysis and J. Amess for haematological diagnosis. Supported by the Imperial Cancer Research Fund.

Received 27 July; accepted 21 September 1987.

1. Rowley, J. D. *Nature* **243**, 290-293 (1973).
2. LeBeau, M. M. & Rowley, J. D. *Cancer Surv.* **3**, 371-384 (1984).
3. Groffen, J. *et al. Cell* **36**, 93-99 (1984).
4. Ben-Neriah, Y., Daley, G. Q., Mes-Masson, A. M., Witte, O. N. & Baltimore, D. *Science* **233**, 212-214 (1986).
5. Clark, S. S., McLaughlin, J., Crist, W. M., Champlin, R. & Witte, O. N. *Science* **235**, 85-88 (1987).
6. Kurzrock, R. *et al. Nature* **325**, 631-635 (1987).
7. Chan, L. C. *et al. Nature* **325**, 635-637 (1987).
8. Third International Workshop on Chromosomes in Leukaemia *Cancer Genet. Cytogenet.* **4**, 95-142 (1981).
9. Rodenhuis, S. *et al. New Engl. J. Med.* **313**, 51-52 (1985).
10. De Klein, A. *et al. Blood* **68**, 1369-1375 (1986).
11. Erikson, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 1807-1811 (1986).
12. Westbrook, C. A. *et al. J. cell. Biochem.* Suppl. **11a**, 200 (1987).
13. Stam, K., Heisterkamp, N., Reynolds, F. H. & Groffen, J. *Molec. cell. Biol.* **7**, 1955-1960 (1987).
14. Hariharan, I. K. & Adams, J. M. *EMBO J.* **6**, 115-119 (1987).
15. Konopka, J. B. *et al. J. Virol.* **51**, 223-232 (1984).
16. Pegoraro, L. *et al. J. natn. Cancer Inst.* **70**, 447-451 (1983).
17. Mes-Masson, A. M., McLaughlin, J., Daley, G. Q., Paskind, M. & Witte, O. N. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9768-9772 (1986).
18. Catovsky, D. *Br. J. Haemat.* **42**, 493-498 (1979).
19. Shtivelman, E. *et al. Cell* **47**, 277-284 (1986).
20. Shtivelman, E. *et al. Blood* **69**, 971-973 (1987).
21. Konopka, J. B. & Witte, O. N. *Biochim. biophys. Acta* **823**, 1-17 (1985).
22. Prywes, R., Hoag, J., Rosenberg, N. & Baltimore, D. *J. Virol.* **54**, 123-130 (1985).
23. Sefton, B. M., Trowbridge, I. S., Cooper, J. A. & Scolnick, E. M. *Cell* **31**, 465-474 (1982).
24. Townsend, A. *et al. Cell* **44**, 959-968 (1986).
25. Rothbard, J., Fernandez, R. & Schoolnik, G. *J. exp. Med.* **160**, 208-221 (1984).
26. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
27. Cooper, J. A. & Hunter, T. *Molec. cell. Biol.* **1**, 165-170 (1981).
28. Th'ng, K. H. *et al. Int. J. Cancer* **39**, 89-93 (1987).

Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes

P. Cantatore*, M. N. Gadaleta*, M. Roberti*, C. Saccone* & A. C. Wilson†

* CSMME del CNR e Dipartimento di Biochimica e Biologia Molecolare, Universita di Bari, 70126 Bari, Italy
† Department of Biochemistry, University of California, Berkeley, California 94720, USA

During the evolution of sea urchins, a transfer RNA gene lost its tRNA function and became part of a protein-coding gene. This functional loss of a tRNA with specificity for one group of leucine codons (CUN, where N is any base) was accompanied by the gain of a new tRNA with that specificity. The new tRNA gene for CUN codons appears to have evolved by duplication and divergence from a tRNA gene specific for another group of leucine codons (UUR, where R is a purine). These proposals account for (1) the strong sequence resemblance between the modern tRNA genes for CUN and UUR codons in *Paracentrotus*, (2) the altered location of the *CUN* gene in mitochondrial DNA of this urchin, and (3) the persistence of a 72-base pair sequence containing a trace of the old *CUN* gene at its original location. The old *CUN* gene now codes for an extra 24 amino acids at the amino end of subunit 5 in NADH dehydrogenase. Besides giving clues about the mechanisms by which tRNA genes move during mitochondrial DNA evolution, this finding leads us to propose a pathway relating the arrangements of other genes in mitochondrial DNAs from four animal phyla.

The sea urchin, *Paracentrotus lividus*, has 22 tRNA genes in mitochondrial (mt) DNA (refs 1, 2) as do the mtDNAs of all other metazoan phyla tested[3,4]. In every case there are two genes for leucine, two for serine, and one each for the remaining 18 amino acids. However, the arrangement of the tRNA genes differs greatly from phylum to phylum. In several phyla, tRNA genes are scattered among genes coding for proteins and may act as punctuation marks in primary transcripts. In sea urchins, by contrast, 15 of the tRNA genes are in a cluster that could have a role in the initiation and regulation of replication and transcription[2]. To throw light on the possible mechanisms of mtDNA rearrangement we have sequenced 20 of the urchin tRNA genes (Cantatore *et al.*, in preparation). As shown in Fig. 1, which presents the base sequences of the two leucine tRNA genes, one specifying a tRNA recognizing CUN codons and the other specifying a UUR tRNA, the two genes are very alike. In pairwise comparisons of 20 sea urchin tRNAs (results not shown), only for the two leucine tRNAs is there as much as an 80% similarity; for all other pairs the percentage similarity lies between 58% and 24%. There are 13 positions at which base differences exist between the two urchin leucine genes (Table 1), whereas the mean number of differences between the urchin *CUN* gene and the *CUN* genes of vertebrates is 36. The low extent of divergence between the two urchin leucine tRNA genes is statistically significant. By applying a Monte Carlo simulation method that takes account of the base compositions of the sequences compared[5], we find that the most probable number of base substitutions (16.7) between the CUN and UUR leucine tRNAs of *Paracentrotus* is more than two standard deviations lower than between any other CUN-UUR couple (Table 1). Moreover, according to tree analysis, both the sea urchin *CUN* and *UUR* genes are related to the *UUR* genes of vertebrate mtDNA (Fig. 2). These observations imply that the *CUN* gene of urchins arose by duplication, and subsequent divergence, from the urchin *UUR* gene.

The approximate time at which the duplication event occurred may be estimated as follows. The most probable number of substitutions between the two urchin genes, 16.7, is similar to that between the leucine tRNA genes of rodents and the cow, namely an average of 17.4 for the *UUR* gene and 15.5 for the *CUN* gene (Table 1). If the rate of substitution has been as high as in mammals, the divergence time between the two urchin genes would be roughly 80 million years. If, however, the rate of evolutionary substitution in urchin mtDNA is lower than in mammals by a factor of two to four[6] (S. Palumbi and A. Wilson, unpublished observations), the divergence may have taken place as early as 320 million years ago. Both times lie within the 500 million year period of diversification of the lineages leading to modern representatives of the echinoderms[7].

The replacement of one *CUN* gene by another that differs greatly in sequence from the first one (presumably at about 36 positions) draws attention to the enormous tolerance for structural variation in animal mitochondrial tRNAs[2,8,9], and strengthens the idea that protein synthesis in animal mitochondria might not be as tightly constrained by demands for high accuracy as in other systems[10,11]. In addition, this late duplication of a tRNA gene, followed by an anticodon change, provides an example of the type of event that could underlie shifts in the genetic code, which have occurred often during mtDNA evolution[1-3,10].

Although the old *CUN* gene has lost its tRNA function, it

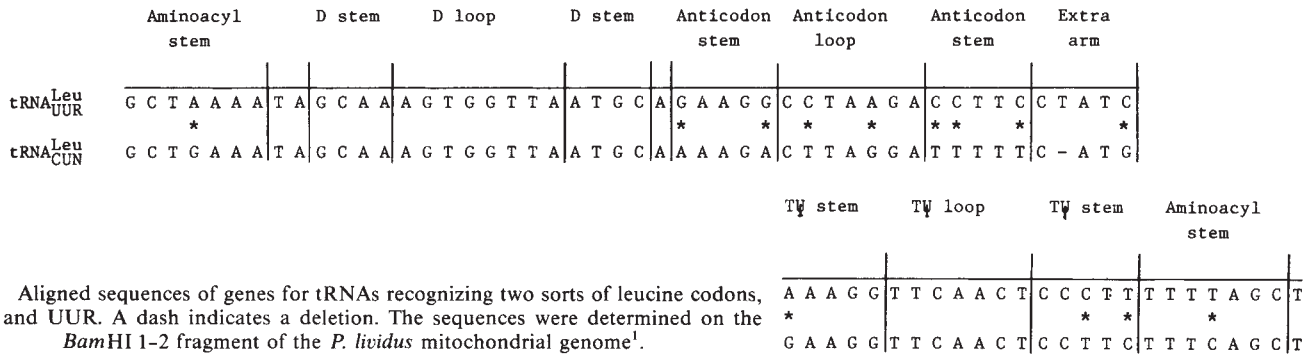


Fig. 1 Aligned sequences of genes for tRNAs recognizing two sorts of leucine codons, CUN and UUR. A dash indicates a deletion. The sequences were determined on the *Bam*HI 1-2 fragment of the *P. lividus* mitochondrial genome[1].

seems to have survived physically in *Paracentrotus* mtDNA and to have acquired a new function. This became evident when defining the structure of the region between the genes coding for the proteins ND4 and ND5, which are subunits of NADH dehydrogenase and designated as *4* and *5* in Fig. 3. In vertebrate mtDNA (refs 3, 4) this region is made up of three tRNA genes, with anticodons for histidine (H), serine (S) and leucine CUN codons (L), the order of the genes being *4, H, S, L, 5*. The sea urchin studied has only two tRNA genes in this region[2], the arrangement being *4, H, S, L*-5*, where *L** stands not for a tRNA gene, but for a 72-base pair (bp) extension of the reading frame of *ND5*. (5′) ATG.GTA.ATA.AAC.CCA.TCA.ACT.ATA .GTA.TCT.ACA.CTT.AGA.ATG.AGA.ATA.ATG.GCT.ATT .CTT.GCA.GTG.AGT.ATA(3′). This 5′ extension makes the urchin ND5 protein 24 amino acids longer at the amino end than its counterpart in vertebrates[8,12,13]. Besides being similar in size to a tRNA gene for leucine, *L** contains a 7-bp sequence (CTTAGAA) near the middle which differs by only one mismatch from the anticodon loop of authentic *CUN* genes. Elsewhere, there is little homology between *L** and vertebrate *CUN* genes, as would be expected for a sequence that diverged about 600 million years ago by both drift and positive selection for a new function. This trace of the anticodon loop bolsters the idea that the 72-bp element at the 5′ end of the *ND5* gene was once the tRNA gene for leucine CUN codons.

The locations, as well as the sequences, of the *UUR* and new *CUN* genes in urchin mtDNA fit the hypothesis that the latter gene arose by duplication of the *UUR* gene rather than of the old *CUN* gene. The urchin *UUR* gene occupies the same place as its counterpart in vertebrate mtDNA next to the 5′ end of the gene coding for the protein ND1 (which is abbreviated as *1* in Fig. 3). The new urchin *CUN* gene is near the *UUR* gene; in fact both reside in the cluster of 15 tRNA genes surrounding the noncoding region (*n* in Fig. 3), which is far from *L**, the old *CUN* gene. The new *CUN* gene may thus have arisen in the vicinity of its source, the *UUR* gene.

It is puzzling that the urchin *UUR* and new *CUN* genes are separated by 8 other tRNA genes. Since the simplest mechanism of gaining a tRNA gene would produce tandem duplicates, another mechanism might be involved, perhaps one that would also account for the generation of additional tRNA genes in

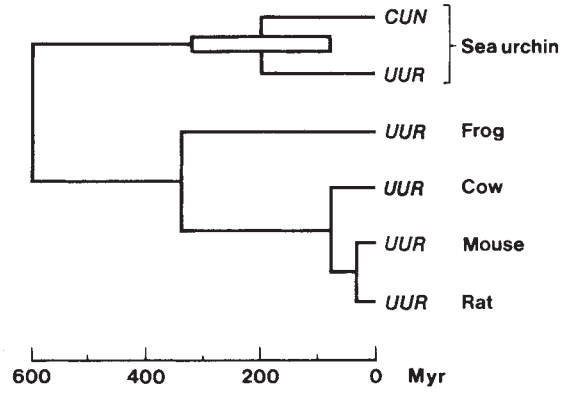


Fig. 2 Tree and tentative time scale for leucine tRNA genes of sea urchin and vertebrate mtDNA. The order of branching was confirmed by parsimony and distance methods applied to the observed number of base differences (Table 1). The time scale is based on the assumption that divergence times are as follows: 600 million years for urchins and vertebrates, 340 million years for frogs and mammals, 80 million years for rodents and cow, 35 million years for rat and mouse. The duplication event that gave rise to the *Paracentrotus* CUN gene may have occurred 80-320 million years ago. A consideration of transition-transversion ratios[10] favours the younger date.

Table 1 Differences between genes coding for CUN and UUR leucine tRNAs

	P_{uur}	M_{uur}	R_{uur}	B_{uur}	X_{uur}	P_{cun}	M_{cun}	R_{cun}	B_{cun}	X_{cun}
P_{uur}	—	29	30	27	21	13	34	34	33	36
M_{uur}	31.0±5.4	—	5	18	21	37	40	41	43	47
R_{uur}	32.1±5.3	10.0±4.3	—	17	24	35	34	38	33	35
B_{uur}	28.9±5.3	20.7±4.8	13.8±4.8	—	21	26	40	42	45	45
X_{uur}	24.1±5.2	27.1±5.2	26.4±5.1	23.7±5.3	—	26	41	40	40	40
P_{cun}	16.7±4.9	37.2±5.5	36.6±5.5	28.2±5.3	28.5±5.5	—	34	38	33	35
M_{cun}	35.3±5.4	40.5±5.4	38.6±5.3	41.2±5.1	41.1±5.4	35.4±5.3	—	4	7	16
R_{cun}	35.5±5.3	42.1±5.4	37.3±5.6	42.9±5.3	40.4±5.0	35.5±5.4	8.9±4.1	—	5	16
B_{cun}	34.7±5.4	43.7±5.2	39.2±5.3	43.8±5.2	40.6±5.3	34.9±5.4	11.5±4.4	19.4±3.9	—	14
X_{cun}	37.4±5.2	47.6±4.9	42.9±5.2	45.6±5.2	41.2±5.1	36.1±5.2	17.2±5.0	19.5±5.0	17.3±4.5	—

Differences due to base substitutions are shown between the genes for the CUN and the UUR leucine tRNAs of sea urchin, frog[12] and mammals[8,9,13]. P stands for *Paracentrotus* (urchin), M for *Mus* (mouse), R for *Rattus* (rat), B for *Bos* (cow), and X for *Xenopus* (frog). The upper right half gives the observed number of base differences. The lower left half contains estimates of the probable numbers of evolutionary substitutions required to account for the observed differences; both the mean values and statistical fluctuations were determined by generating, with a Monte Carlo method described by Lanave *et al.*[5], about 5,000 sequences of fixed length.

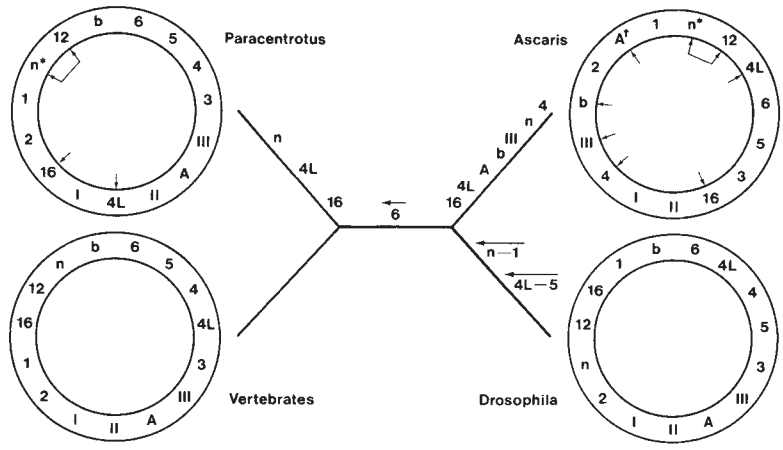


Fig. 3 Unrooted tree giving one possible explanation for the differences in gene order found among mtDNAs of four animal phyla[1,3,4,15,17]. The 15 genetic regions considered are abbreviated as follows: *1, 2, 3, 4, 4L, 5* and *6* (subunits of NADH dehydrogenase), *12* and *16* (small and large ribosomal RNAs), *I, II* and *III* (cytochrome oxidase subunits), *b* (cytochrome b), *A* (ATPase subunits) and *n* (main noncoding region). All tRNA genes are excluded from this analysis. Above each segment of the tree, the genes that moved are listed; horizontal arrows indicate polarity changes. Within the circular maps, arrows point to genes that moved, the asterisk indicates that the event affecting the relation of regions *12* and *n* may not be the same in *Ascaris*[17] and *Paracentrotus*[1], and the dagger indicates that *A* moved and lost subunit 8.

this region. It has recently been shown[14] that, in the replication origin for the heavy strand of mammalian mtDNA, the nucleotide sequence can be folded in tRNA-like structures. The replication origin of the mammalian light strand is also surrounded by tRNA genes. Hence, tRNA-like structures, probably acting as primers, may be necessary for DNA synthesis. So, we ask whether the appearance of tRNA genes at new locations could be mediated by tRNAs. If tRNAs could themselves act as primers for mtDNA synthesis and sometimes fail to be removed from the newly synthesized strand, the tRNA would become a template in the next round of replication. The result would be the incorporation of a tRNA gene at the replication origin.

Whether or not this hypothetical mechanism exists, the urchin findings reinforce the view[15] that tRNA genes can move independently of other mitochondrial genes. Given this assumption, it becomes easier to explain the differences between phyla in the order of other genes in mtDNA. Fig. 3 shows the orders of the genes for proteins and ribosomal RNAs in representatives of four phyla. Only a few intergenic recombination events are needed to interrelate these gene orders. There appear to have been no rearrangements on the vertebrate lineage, which implies that the urchin arrangement arose from the vertebrate one. The absence of rearrangements during vertebrate evolution fits with other evidence indicating that vertebrate mitochondria could lack the capacity to carry out intergenic recombination[16]. Tests of these proposals, based on Fig. 3, can be made experimentally and by defining gene orders for additional species. Such information might also contribute to knowledge of relationships among the primary branches in the genealogical tree for metazoa.

We thank C. Lanave for helping with the Monte Carlo analysis and D. Wolstenholme and G. Dover for comments on the manuscript. This work received support from Progetto Finalizzato Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie e Progetto Strategico Biotecnologie, CNR, Italy and from MPI. The paper was written while A. C. Wilson was on sabbatical leave in Bari and at the Genetics Department, Cambridge University.

Received 20 July; accepted 18 August 1987.

1. Cantatore, P. *et al. Gene* **53**, 41-54 (1987).
2. Cantatore, P., Roberti, M., Rainaldi, G., Saccone, C. & Gadaleta, M. N. *Curr. Genet.* (in the press).
3. Attardi, G. *Int. Rev. Cytol.* **93**, 93-145 (1985).
4. Cantatore, P. & Saccone, C. *Int. Rev. Cytol.* **108**, 149-208 (1987).
5. Lanave, C., Preparata, G., Saccone, C. & Serio, G. *J. molec. Evol.* **20**, 86-93 (1984).
6. Vawter, L. & Brown, W. M. *Science* **234**, 194-196 (1986).
7. Smith, A. *Echinoid Palaeobiology* (Allen & Unwin, London, 1984).
8. Bibb, M. J., Van Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. *Cell* **26**, 167-180 (1981).
9. Cantatore, P. *et al. Nucleic Acids Res.* **10**, 3279-3289 (1982).
10. Wilson, A. C. *et al. Biol. J. Linn. Soc. Lond.* **26**, 375-400 (1985).
11. Pietromonaco, S. F., Hessler, R. A. & O'Brien, T. W. *J. molec. Evol.* **24**, 110-117 (1986).
12. Roe, B. A., Ma, D.-P., Wilson, R. K. & Wong, J. F.-H. *J. biol. Chem.* **260**, 9759-9774 (1985).
13. Anderson, S. *et al. J. molec. Biol.* **156**, 683-717 (1982).
14. Brown, G. G., Gadaleta, G., Pepe, G., Saccone, C. & Sbisà, E. *J. molec. Biol.* **192**, 503-511 (1986).
15. Clary, D. O. & Wolstenholme, D. R. *J. molec. Evol.* **22**, 252-271 (1985).
16. Olivo, P. D., Van den Walle, M. J., Laipis, P. J. & Hauswirth, W. W. *Nature* **306**, 400-402 (1983).
17. Wolstenholme, D. R., Macfarlane, J. L., Okimoto, R., Clary, D. O. & Wahleithner, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1324-1328 (1987).

Identification of the V factor needed for synthesis of the iron-molybdenum cofactor of nitrogenase as homocitrate

Timothy R. Hoover*†, Andrew D. Robertson*, Ronald L. Cerny‡, Roger N. Hayes‡, Juan Imperial*†, Vinod K. Shah*† & Paul W. Ludden*†§

* Department of Biochemistry and † Center for the Study of Nitrogen Fixation, College of Agriculture and Life Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA
‡ Midwest Center for Mass Spectrometry, Department of Chemistry, University of Nebraska-Lincoln, Lincoln, Nebraska 68588, USA

Nitrogenase catalyses the ATP-dependent reduction of N_2 to NH_3, and is composed of two proteins, dinitrogenase (MoFe protein or component I) and dinitrogenase reductase (Fe protein or component II)[1,2]. Dinitrogenase contains a unique prosthetic group (iron-molybdenum cofactor, FeMoco) comprised of Fe, Mo and S, which has been proposed as the site of N_2 reduction[3–5]. Biochemical and genetic studies of Nif⁻ (nitrogen fixation) mutants of *Klebsiella pneumoniae* which are defective in nitrogen fixation, have shown that the *nifB*, *nifQ*, *nifN*, *nifE* and *nifV* genes are required for the biosynthesis of FeMo-co[5–7]. Recently, a system for *in vitro* synthesis of FeMoco was described[8]. The assay requires at least the *nifB*, *nifN* and *nifE* gene products[8], and a low-molecular-weight factor (V factor) produced in the presence of the *nifV* gene product[9]. We have used this system to study FeMoco biosynthesis. We report here the isolation of V factor and identify it as homocitric acid ([*R*]2-hydroxy-1,2,4-butanetricarboxylic acid).

Initial attempts to purify V factor used extracts of *K. pneumoniae*[9]. V factor was found to accumulate in the medium of cultures of *K. pneumoniae* which had been derepressed for nitrogenase synthesis, and V factor was subsequently isolated from the medium. Cultures of wild-type *K. pneumoniae* (strain UN) excreted V factor into the medium, whereas cultures of a *nifV* mutant (UN1991) did not (Fig. 1). The amount of V factor in the medium of the wild-type culture continued to increase

§ To whom correspondence should be addressed.

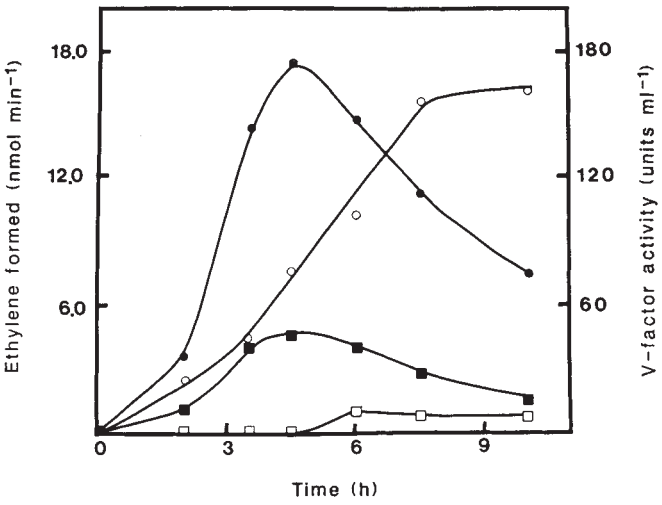


Fig. 1 Accumulation of V factor in the medium during derepression of nitrogenase. Cultures of *K. pneumoniae* were derepressed for nitrogenase as described previously[7]. Whole-cell acetylene reduction activities (filled symbols) were determined for cultures of ●, UN a reisolate of the wild-type strain M5a1; and ■, UN1991 (*nifV4945*)[15]. V factor (open symbols) was assayed as described[9], with modifications noted below, from 10 and 25 μl of medium from cultures of ○, UN and □, UN1991.
Methods. For the V-factor assay, 100 μl of preparation containing the partially purified *nifB* gene product from *K. pneumoniae* UN1100 (*nifE4420*) (1 mg protein) were mixed with 200 μl of desalted *Azotobacter vinelandii* UW45 extract (4 mg protein), 200 μl of an ATP-generating system containing 50 μM Na_2MoO_4 and 5 mM sodium dithionite, and 100 μl of anaerobic Tris-HCl buffer. The reaction mixture was incubated for 30 min at 30 °C. An additional 800 μl of ATP-generating mix containing 5 mM sodium dithionite, along with 30 μg of purified dinitrogenase reductase from *A. vinelandii* were added to the reaction mixture, and acetylene reduction assays were done as described[9]. Before the assay for V factor, media samples were cleared by centrifugation for 1 min with a microcentrifuge. One unit of V factor corresponds to the amount required to give a twofold increase in the activity in the absence of added V factor, as measured by the coupled acetylene reduction assay.

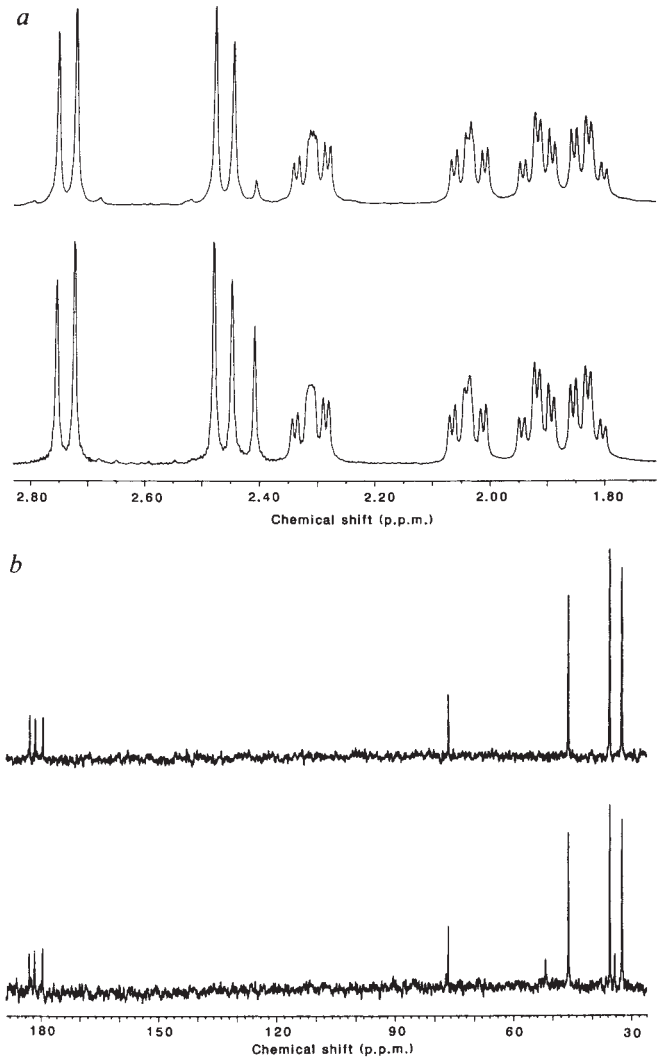


Fig. 2 The ^1H (*a*) and ^{13}C (*b*) NMR spectra of homocitrate (top traces) and V factor (bottom traces) were compared.
Methods. The sample (8 mg) was dissolved in 0.5 ml ^2H_2O and the *p*D adjusted to 11.9. TSP (3-(trimethylsily) propionic acid-2,2,3,3-^2H) (10 μmol) was added as an internal ^1H chemical shift standard. Dioxane in ^2H_2O (40%, v/v) was used as an external ^{13}C chemical shift standard. All NMR spectra were obtained on a Bruker AM 500 spectrometer operating at 500.13 MHz ^1H frequency, 125.77 MHz ^{13}C frequency, and equipped with an Aspect 3000 computer and digital phase shifters. Sample temperature was maintained at 25 °C throughout all experiments. The carrier frequency was always set in the middle of the spectrum. One-dimensional ^1H NMR spectra were obtained with a single-pulse experiment with a total sweep width of 3,100 Hz. A total of 8,192 time -domain data points were acquired during the 1.32-s acquisition time; eight scans were averaged. The data were multiplied by an exponential decay function, with a line-broadening parameter of 1.5 Hz, before Fourier transformation. The ^1H-decoupled ^{13}C NMR data were acquired with composite pulse decoupling[16]. The sweep width was 25,000 Hz and between 650 and 750 scans, each consisting of 16,384 time-domain data points, were averaged in each experiment. The data were multiplied by an exponential decay function, with a line-broadening parameter of 15 Hz, before Fourier transformation.

even after the peak of whole-cell C_2H_2 reduction activity. V factor accumulated in the media of cultures of all Nif mutants examined, except for NifV⁻ mutants (ignoring NifA⁻ mutants, which lack the protein required for the expression of all the other *nif* genes). These data indicate that the *nifV* gene product is not merely involved in the uptake of V factor, but is required for the synthesis of V factor.

The structure of V factor was determined by NMR spectroscopy and mass spectrometry (MS). With ^1H-^1H homonuclear correlation, direct ^{13}C-^1H heteronuclear correlation and long-range ^{13}C-^1H heteronuclear correlation experiments, it was possible to establish homocitrate as a model compound for the structure of V factor. Confirmation that V factor is homocitrate is shown in Fig. 2, as the ^1H and ^{13}C NMR spectra of V factor and homocitrate (Sigma) are identical. The only observed contaminant (data not shown for ^1H NMR spectrum) in the V factor preparation was identified as tetraethylammonium chloride, an HPLC ion-pairing reagent used in the isolation procedure. The NMR spectra were more complicated at low *p*D owing to the formation of a stable five-membered lactone ring involving the hydroxyl group and the 6-carboxyl group.

Positive-ion fast atom bombardment (FAB) MS was used as further confirmation that V factor is homocitrate. High-resolution FAB MS experiments gave observed masses of 189.0395 mass units for the $(M+H)^+$ species of V factor (lactone), and 189.0396 mass units for the $(M+H)^+$ of homocitric acid lactone. The theoretical mass of $C_7H_9O_6$ is 189.0399 mass units. The FAB MS-MS spectra of the $(M+H)^+$ species ($m/z=189$) of both V factor and homocitric acid lactone had significant peaks of $m/z=171$ $(M+H-H_2O)^+$, 143 $(M+H-HCO_2H)^+$, and 115 $(M+H-HCO_2H-CO)^+$.

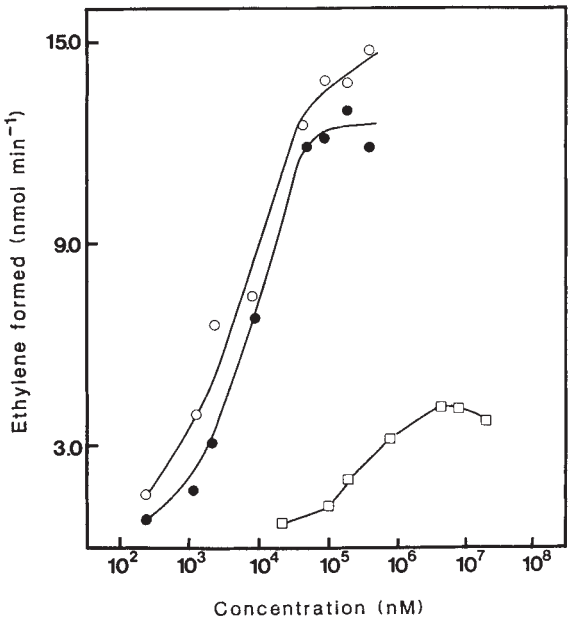


Fig. 3 Requirement of homocitrate for the *in vitro* synthesis of FeMo-co. The V factor assay mixture was titrated with ○, V factor (1 mg ml^{-1}); ●, homocitrate (5 mM); and □, citrate (0.5 M). V factor and homocitrate were dissolved in 50 mM NaOH to prevent formation of the lactone. Concentrations are given for the 30-min incubation before acetylene reduction assays. The molecular weight of homocitrate was used to calculate the concentration of V factor added to the assay.

Methods. V factor was purified from the culture medium of UN1100 as follows. Medium (14 l) was loaded onto a 32×5 cm AG 1-X8 (Bio-Rad) anion exchange column. The column was washed with H_2O, followed by 0.35 M NaCl and 0.40 M NaCl (which contained the V factor). The *p*H of the 0.40 M NaCl fraction was adjusted to 2.0 with 4 M HCl. The fraction was dried under reduced pressure and V factor was extracted with acetone. The acetone extract was vacuum dried, the residue dissolved in H_2O, and the *p*H of the solution adjusted to 5.5 with 4 N NaOH. The sample was loaded onto an hydroxylapatite (Bio-Rad HPT) column which had previously been equilibrated with H_2O. The column was washed with two bed volumes of H_2O, followed by two bed volumes of 0.1 M KH_2PO_4 buffer, *p*H 7.4. Most of the V factor activity was collected in the 0.1 M KH_2PO_4 fraction, which was acidified, vacuum dried, and extracted with acetone. After removing the acetone under reduced pressure, the residue was dissolved in 100 mM formic acid. Final steps of the V factor purification involved the sequential use of a reverse-phase C18 HPLC column (600 RPB; Alltech) with each of the three following solvent systems: 100 mM formic acid; 10 mM tetraethylammonium chloride in 100 mM formic acid; or 0.1% (v/v) trifluoroacetic acid. HPLC fractions with V factor activity were pooled, vacuum dried, and the residue dissolved in the appropriate solvent before being applied to the HPLC column.

Additional proof that V factor is homocitrate is shown in Fig. 3, where homocitrate replaced V factor in the assay for FeMoco synthesis. Addition of homocitrate to the assay resulted in the synthesis of active FeMoco to about the same extent as V factor. The concentrations required to reach the same level of FeMo-co synthesis were about twofold higher for the synthetic homocitrate compared to the natural product isolated from *K. pneumoniae* medium, probably because the synthetic homocitrate was a racemic mixture. These data, together with the NMR spectroscopy and MS data, show unambiguously that V factor is homocitrate.

Several di- and tricarboxylic acids were tested for their ability to replace homocitrate in the assay (data not shown). Only citrate, at concentrations 100-fold higher than those required for homocitrate, showed appreciable activity in the assay (Fig. 3). Mutants affected in *nifV* have been reported to produce FeMoco with altered substrate specificity and inhibitor susceptibility[5]. It will be interesting to determine if the substitution of citrate for homocitrate for the *in vitro* synthesis of FeMoco will mimic these results.

The optical rotation of V factor (lactone) in aqueous solution was measured at 21 °C with a Perkin-Elmer Model 141 Polarimeter at 589 nm. The specific rotation was calculated to be −43.7°. Thomas *et al.*[10] synthesized (*S*)-2-hydroxy-1,2,4-butanetricarboxylic acid and reported specific rotations of +52.6° and +60°. Therefore, the absolute configuration of the natural product from *K. pneumoniae* may be designated as (*R*)-2-hydroxy-1,2,4-butanetricarboxylic acid.

The involvement of homocitrate in FeMoco synthesis represents the first report of the presence and function of this compound in a prokaryotic organism. The only previously known biological role of homocitrate is as an intermediate in lysine biosynthesis (2-aminoadipate pathway) in yeasts and fungi[11,12]. In *Saccharomyces cerevisiae*, homocitrate is formed by the condensation of acetyl-CoA with 2-oxoglutarate[11,13]. The enzyme catalysing this reaction is homocitrate synthase (EC 4.1.3.21)[13,14]. Like the natural product from *K. pneumoniae*, the absolute configuration of the intermediate in lysine biosynthesis is designated as (*R*)-2-hydroxy-1,2,4-butanetricarboxylic acid[10].

Based on our identification of V factor as homocitrate, we propose that the *nifV* gene encodes homocitrate synthase. We are trying to develop an assay for homocitrate synthase from *K. pneumoniae*.

Homocitrate may function in the uptake, initial processing, or targeting of metal ions for FeMoco synthesis. Ultraviolet-visible difference spectra indicate that homocitrate will form complexes with both Fe^{3+} and MoO_4^{2-} (data not shown). Alternatively, homocitrate may be incorporated into FeMoco, but it has not been possible to demonstrate the release of V factor activity from acid-hydrolysed dinitrogenase or O_2-denatured FeMoco[9]. We are designing experiments to determine the fate of homocitrate in the *in vitro* FeMoco synthesis assay. This will provide information on the involvement of homocitrate in FeMoco biosynthesis.

We thank Claire Knocke for technical assistance, Mark Pope and W. Milo Westler for suggestions; Mina Nashed and Laurens Anderson for assistance with polarimetry experiments; Winston Brill for his interest and encouragement, and Gary Roberts for evaluating the manscript. This research was supported by the College of Agriculture and Life Sciences, University of Wisconsin and an NIH grant to P.W.L. and V.K.S. NMR experiments were conducted at the National Magnetic Resonance Facility at Madison which is supported by grants from NIH, the NSF Biological Instrumentation Program and the USDA. The Midwest Center for Mass Spectrometry is supported by an NSF grant. T.R.H. was supported by a Procter and Gamble fellowship. J.I. is on leave from the C.S.I.C. of Spain. P.W.L. is a Harry and Evelyn Steenbock Career Development Award recipient.

Received 14 August; accepted 17 September 1987.

1. Bulen, W. A. & LeComte, J. R. *Proc. natn. Acad. Sci. U.S.A.* **56**, 979-986 (1966).
2. Hageman, R. V. & Burris, R. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2699-2702 (1978).
3. Shah, V. K. & Brill, W. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3249-3253 (1977).
4. Rawlings, J. *et al. J. biol. Chem.* **253**, 1001-1004 (1978).
5. Hawkes, T. R., McLean, P. A. & Smith, B. E. *Biochem J.* **217**, 317-321 (1984).
6. Roberts, G. P., MacNeil, T., MacNeil, D. & Brill, W. J. *J. Bact.* **136**, 267-279 (1978).
7. Imperial, J., Ugalde, R. A., Shah, V. K. & Brill, W. J. *J. Bact.* **163**, 1285-1287 (1985).
8. Shah, V. K., Imperial, J., Ugalde, R. A., Ludden, P. W. & Brill, W. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1636-1640 (1986).
9. Hoover, T. R., Shah, V. K., Roberts, G. P. & Ludden, P. W. *J. Bact.* **167**, 999-1003 (1986).
10. Thomas, U., Kalyanpur, M. G. & Stevens, C. M. *Biochemistry* **5**, 2513-2516 (1966).
11. Strassman, M. & Weinhouse, S. *J. am. chem. Soc.* **75**, 1680-1684 (1953).
12. Broquist, H. & Trupin, J. *A. Rev. Biochem.* **35**, 231-274 (1966).
13. Strassman, M. & Ceci, L. N. *Biochem. Biophys. Res. Commun.* **14**, 262-267 (1964).
14. Tucci, A. F. & Ceci, L. N. *Arch. Biochem. Biophys.* **153**, 742-750 (1972).
15. MacNeil, T., MacNeil, D., Roberts, G. P., Supiano, M. A. & Brill, W. J. *J. Bact.* **136**, 253-266 (1978).
16. Shaka, A. J., Keeler, J., Frankel, T. & Freeman, R. *J. magn. Reson.* **52**, 335-338 (1983).

Distal residues in the oxygen binding site of haemoglobin studied by protein engineering

K. Nagai*, B. Luisi*||, D. Shih*||, G. Miyazaki*||, K. Imai*||, C. Poyart†, A. De Young‡, L. Kwiatkowsky‡, R. W. Noble‡, S.-H. Lin§ & N.-T. Yu§

* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK
† INSERM Unite 299, 78 rue du Général Leclerc, 94274 Le Kremlin-Bicêtre, France
‡ Department of Medicine and Biochemistry, State University of New York at Buffalo, Veterans Administration Medical Center, Buffalo, New York 14215, USA
§ School of Chemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

The geometries of the Fe–O_2 and Fe–CO bonds in myoglobin and haemoglobin differ significantly from those in free porphyrin model compounds[1–6]. It has been suggested that steric hindrance by Val-E11 and His-E7 and a hydrogen bond between His-E7 and oxygen[2,4,7] affect the geometry and electronic state of the Fe-ligand bond, and that these interactions may be important in controlling oxygen affinity[8]. We have produced mutant haemoglobins in *E. coli*[9–11] having Val(67β)E11 replaced by Ala, Met, Leu or Ile and His(58β)E7 by Gln, Val or Gly. We have studied the effect of these mutations on the equilibrium and kinetics of ligand binding. The conformation of the new side chains and their effect on the protein structure have been examined by X-ray crystallography, and the vibrational properties of the Fe–CO bond observed by resonance Raman spectroscopy[12]. We found that the steric hindrance of ligand binding by the E11 residue and the polarity of the E7 residue in the β subunit are critical for fine-tuning ligand affinity.

Oxygen binds to the densely packed interior of haemoglobin (Hb), even though there is no room for it in deoxy-Hb[13–15]. Even the most conservative amino-acid substitution here is accompanied by a volume change comparable to the size of the ligands. By introducing several similar amino acids differing only in size and conformation, it was possible to determine the importance of steric hindrance by Val-E11β. Figure 1*a* shows a difference Fourier map of the Val-E11β substituted Ile mutant minus normal Hb A, at 2.45 Å resolution. The forked side chains of Val and Ile cannot rotate about the C_α-C_β bond when they are attached to an α-helix[16], and so the δ-methyl group of the Ile-E11β side chain is forced to cover the oxygen binding site and hinder oxygen binding. Pairs of positive and negative peaks next to His-E7β and Leu(28β)B10 show that they are pushed away by the Ile-E11β side chain. The difference Fourier map of the Leu-E11β mutant minus Hb A at 2.18-Å resolution shows a large positive peak due to $C_{\delta1}$ and $C_{\delta2}$ of the new Leu side chain and two negative peaks due to loss of $C_{\gamma1}$ and $C_{\gamma2}$ of Val-E11β (Fig. 1*b*).

Unlike the Ile side chain, the Leu side chain can rotate about the C_α-C_β and C_β-C_γ bonds and is therefore more flexible. The Leu-E11β side chain does not extend to the oxygen binding site and causes no movement of His-E7β, but does seem to cause a small movement of the E-helix. The Leu-E11β side chain must be disordered, because we could not fit it into the ($2F_{mutant}$ minus F_{native}) map with a single conformation.

In the crystal structure of Hb A, there are no channels to let the ligands into the haem pocket[13,15]. The substitution of His-E7β by Gly leaves a large space in the oxygen binding site (Fig. 1*c*). When the haem is seen from outside, there is now no obstruction for ligand entry and this mutant binds CO ten times faster than Hb A. The His-E7β to Gly substitution causes surprisingly small structural changes in the molecule. Despite the large gap, there is no significant movement of the E-helix or of the haem. Figure 1*d* shows a refined ($2F_{observed}$ minus $F_{calculated}$ map of the Gln-E7β mutant at 2.12-Å resolution. We cannot distinguish between N_ε and O_ε of the Gln-E7β side-chain at this resolution, but N_ε of Gln-E7β can occupy the position where N_ε of His-E7β is normally found. Gln-E7β is well localized; a water molecule is hydrogen-bonded to either N_ε or O_ε. The Gln side-chain is more flexible than the His side-chain and can probably donate a hydrogen bond to the oxygen molecule with better geometry. The oxygen equilibrium curve of the Leu-E11β mutant is almost identical to that of Hb A (Fig. 2). Owing to its flexibility, the side chain of Leu-E11β exercises no additional steric effect over the normal Val side-chain. The oxygen affinity of the Ile-E11β mutant is lower than that of Hb A. If we apply the generalized allosteric model[20], which allows the oxygen affinity of the α and β subunits to be different, then K_1 and K_4 are given by the following equations[21]:

$$K_1 = (1/2)(K_{T\alpha} + K_{T\beta})$$

and

$$K_4 = \frac{2K_{R\alpha}K_{R\beta}}{K_{R\alpha} + K_{R\beta}}$$

where $K_{R\alpha}$ and $K_{T\beta}$ are the oxygen association constants of the α and β subunits in the T state, and $K_{R\alpha}$ and $K_{R\beta}$ are the oxygen association constants of the α and β subunits in the R state[21]. We assume that this mutation only affects $K_{T\beta}$ and $K_{R\beta}$, and that in Hb A the α and β subunits are equivalent in both the R and T states. With these assumptions, even if the affinity of the β subunit in the T state were practically zero, K_1 would only be halved. In fact, the K_1 value for the Ile-E11β mutant is found to be almost halved, which suggests that $K_{T\beta}$ is substantially reduced. K_4 is lowered about three-fold and $K_{R\beta}$ is about one-fifth of the normal K_R value. The steric hindrance by Val-E11β to oxygen binding is relieved by its replacement with Ala, which increases $K_{T\beta}$ five-fold, but does not influence K_R, showing that the steric effect of the Val-E11β side-chain is significant only in the T state. K_1 is similarly increased by the Gln-E7β mutation. The Gly-E7β and Val-E7β mutants are unstable, which prevented estimates of K_1 and K_4.

We have studied the effects of these amino-acid replacements on the Fe-ligand vibrational modes (Fig. 3). The Raman lines of carbon monoxy haemoglobin (CO-Hb) at 507 cm^{-1}, 578 cm^{-1} and 1951 cm^{-1} have been assigned to the Fe–CO stretching, Fe–C–O bending and C–O stretching modes[12]. The structure of the Ile-E11β mutant shows that the δ-methyl group of Ile-E11β imposes additional steric hindrance on the Fe-ligand bond, but without affecting these vibrational modes. They also remained unaffected by the Val-E11β → Ala, Leu or Met mutations, suggesting that the globin structure in the R state is flexible enough to accommodate these mutations without significantly distorting the Fe–C–O bond. On the other hand, the CO vibrational modes are affected by the replacements of His-E7β with non-polar amino acids. The His-E7β substitution with Val mutant shows a broad shoulder on the low frequency side of the Fe–CO stretching line and the C–O stretching line at 1951 cm^{-1} splits into two lines, one at 1951 cm^{-1} and the other at 1971 cm^{-1}. The mutation is unlikely to affect the α-haem and therefore it must be the β subunit that shows a decrease in the Fe–CO stretching frequency and an increase in the C–O stretching frequency. The replacement of His-E7β with the smaller Gly residue has a similar effect, but the C–O stretching line of the β subunits splits further into two lines, one at 1,965 cm^{-1} and the other at 1,971 cm^{-1}. Multiple C–O stretching bands have been observed in carbon monoxy-myoglobin (CO-Mb) (refs 12, 23). In contrast, the resonance Raman spectrum of the Gln-E7β mutant is iden-

|| Present addresses: Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA (B.L.); Department of Biochemistry, Oregon Health Sciences University, Portland, Oregon 97201, USA (D.S.); Department of Biophysics, Faculty of Engineering Science, Osaka University, Toyonaka-560, Japan (G.M.); Department of Physicochemical Physiology, Osaka University Medical School, Osaka-530, Japan (K.I.).

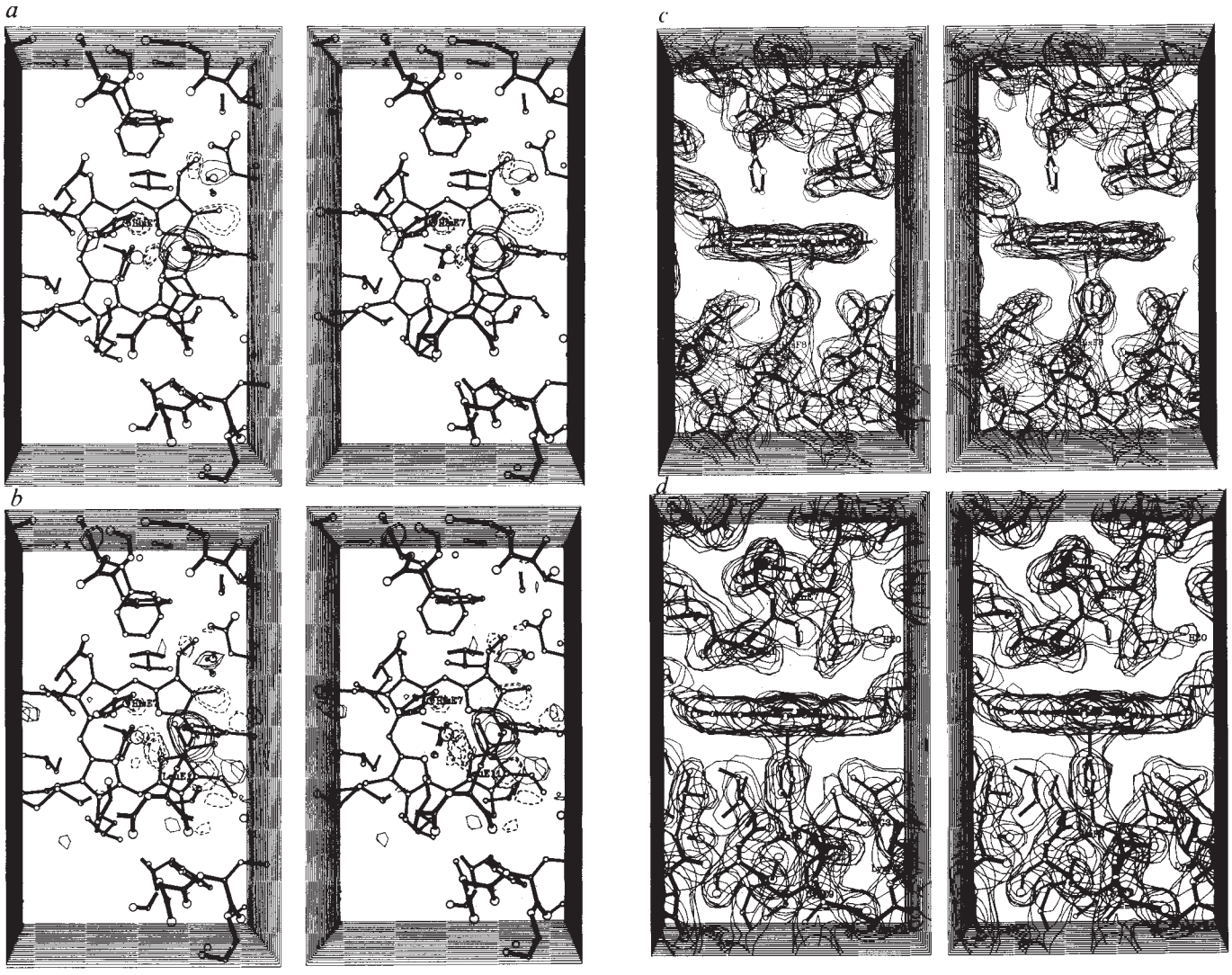


Fig. 1 Structures of mutant Hbs produced in *E. coli*. *a*, A difference Fourier map of the Ile-E11β mutant at 2.45 Å resolution. Solid contours, 0.09 e Å^{-3}; broken contours, −0.09 e Å^{-3}. *b*, A difference Fourier map of the Leu-E11β mutant at 2.18 Å resolution. Solid contours, 0.09 eÅ^{-3}; broken contours, −0.09 e Å^{-3}. *c*, A ($2F_{\text{observed,mutant}} - F_{\text{observed,native}}$) map of the Gly-E7$\beta$ mutant at 2.8 Å resolution. The electron density map contoured at 0.3 e Å^{-3} is superimposed on the atomic model of HbA. *d*, A ($2F_{\text{observed,mutant}} - F_{\text{calculated,mutant}}$) map of the Gln-E7$\beta$ mutant at 2.18 Å resolution. The electron density map contoured at 0.25 eÅ^{-3} is superimposed on the refined atomic model of the mutant.

Methods. Deoxy-Hb crystals were prepared as described[17]. All the mutant crystals were isomorphous with native deoxy-Hb A $P2_1$ crystals. Complete datasets of unique reflections (F_+) were collected at room temperature on a Hilger-Watts 4-circle or a Nonius CAD4 diffractometer. The data were corrected for Lorentz and polarization effects and a linear degradation correction was applied. The datasets of mutant Hbs were scaled to the F_+ from the 1.74 Å native deoxy-Hb A dataset of Fermi *et al.*[13] with *R* factors of 8.8% (Ile-E11β mutant), 9.9% (Leu-E11β), 16.1% (Gly-E7β) and 11.1% (Gln-E7β). The difference density maps were drawn at the contour level of $\pm 3\sigma$ where σ is calculated as the mean difference density over the whole asymmetric unit. The structure of the Gln-E7β mutant was refined with 2 cycles of the Jack-Levitt procedure[18] using as coefficients $F_{\text{diff}} = E(+)_{\text{obs,mut}} - F(+)_{\text{obs,mat}} + F_{\text{cal,nat}}$, where obs, cal, mut and nat denote observed, calculated, mutant and native, respectively[19]. An initial model of the Gln-E7β residue was prepared by manually fitting a Gln side chain using the $2F(+)_{\text{obs,mut}} - F(+)_{\text{obs,nat}}$ map. Shifts were applied only to the main chain and side chains of the residues E5 to E9 of both β subunits. $F_{\text{cal,nat}}$ includes contribution for water molecule and bulk solvent scatter. Temperature factors were taken from the averaged value for the atoms of native His-E7β (ref. 13) and were not refined. The shifts on the Gln atoms were small in the second refinement cycle, less than 0.15 Å. A bulk solvent correction was added to the Gln-E7β model after refinement, which scaled to 0.34 eÅ^{-3}, compared to 0.38 e Å^{-3} for native[13]. The final *R* factor for F_{diff} to $F_{\text{cal,mut}}$ was 13.2%, and the F_{diff} minus $F_{\text{cal,mut}}$ difference map had minimal, maximal and root-mean-square density of −0.83, 0.34 and 0.05 e Å^{-3}.

tical to that of Hb A. As shown by our X-ray analysis, N_ε of Gln-E7β can interact with the ligand in the same way as N_ε of His-E7β. Although a neutron diffraction study of CO-Mb showed no hydrogen bond between CO and the distal His-E7 (ref. 24), polarity of the E7β residue affected the nature of the Fe-CO bond.

Although the Ile and Leu side-chains have the same volume, the functional properties of these two mutants are distinctly different. The δ-methyl group of the Ile E11β blocks the oxygen-binding site and this steric hindrance cannot be relieved because rotation of the Ile side chain is prevented by the α-helix. Hence the Ile-E11β mutant has a lower oxygen affinity than Hb A, and reacts with CO more slowly. By contrast, the Leu-E11β side-chain is flexible and disordered. It does not extend to the oxygen binding site and the oxygen affinity is therefore normal. On the other hand, the replacement of Val-E11β by the smaller Ala residue caused an increase in the oxygen affinity of the T state and the rate of CO combination. The rate of CO combination to the E7β and E11β mutants shows very good correlation with the size the side chain. From these results, we conclude

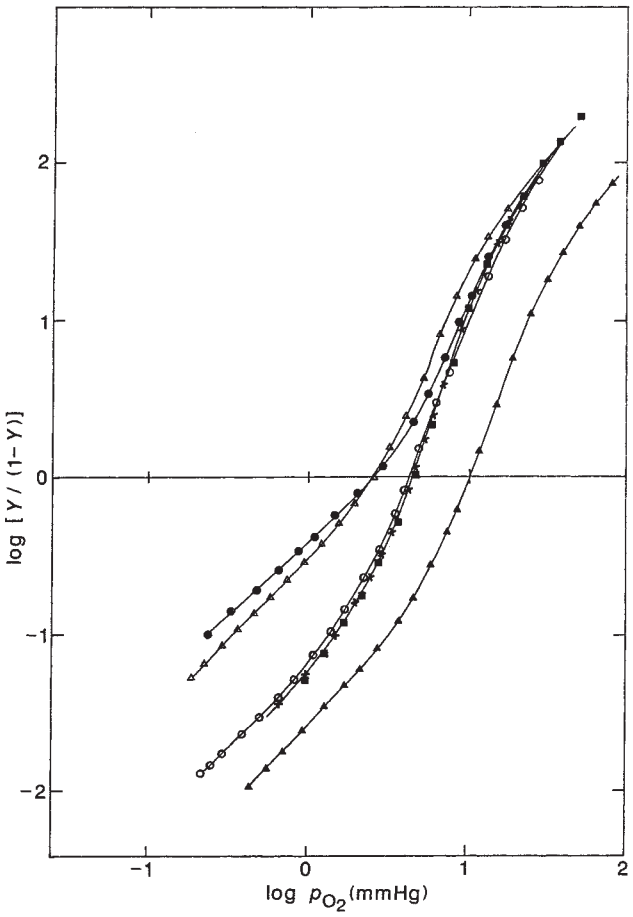


Fig. 2 The Hill plots of oxygen equilibrium curves of mutant Hbs. ●, $\alpha_2\beta$(Val-E11→Ala)$_2$; △, $\alpha_2\beta$(His-E7→Gln)$_2$; ○, $\alpha_2\beta$(Val-E11→Leu); ▲, $\alpha_2\beta$(Val-E11→Ile)$_2$; ×, human Hb A ($\alpha_2\beta_2$); ■, wild-type Hb A produced in *E. coli*. *Y* denotes the fractional saturation of Hb with oxygen. Each curve tends to a straight line of slope unity at each extremity. The intercept of the lower asymptote on the log *p* axis gives $-\log K_1$ and that of the top asymptote $-\log K_4$ where K_1 and K_4 are the oxygen equilibrium constants for the first and fourth oxygen molecules. When the oxygen equilibrium curve is highly cooperative, K_1 and K_4 represent the oxygen affinity of the T and R quaternary structures (K_T and K_R) (ref. 20).

Methods. Mutant β globins were produced in *E. coli* and reconstituted to $\alpha_2\beta_2$ tetramers as described previously[9,10], but purified further on a DEAE-Serphacel column (Shih and Nagai, unpublished results). Note that the oxygen equilibrium curves of Hb A produced in *E. coli* were indistinguishable from those of native human Hb A after further purification. We found previously[10] that the Hill's coefficient of Hb A produced in *E. coli* was lower than that of native human Hb A. Oxygen equilibrium curves were determined in 0.05 M bis-Tris, 0.1 M Cl^{-1} buffer, *p*H 7.4 (ref. 22). Methaemoglobin concentrations recorded after taking measurements were less than 5%, as estimated by optical density at 630 nm.

that the steric hindrance by Val-E11β is important in regulating the oxygen affinity in the T state.

The conclusion is well supported by the amino-acid substitutions observed in evolution. Val-E11 is replaced by Ile in only one out of about one hundred species in the β subunit and eight in the α subunit. The side chain of Val-E11 lies further away from the ligand binding site in the α subunit and therefore Ile is more acceptable there[8,13]. The properties of the Leu-E11β mutant are more similar to those of Hb A and fail to explain why Val-E11 is not replaced by Leu in nature. The flexible side chain of Leu-E11β probably protects the haem iron against oxidation less. His-E7 is even more conserved and is substituted by Gln only in the α subunit of opposum Hb (ref. 25). Our crystallographic study shows that His-E7β and Gln-E7β are

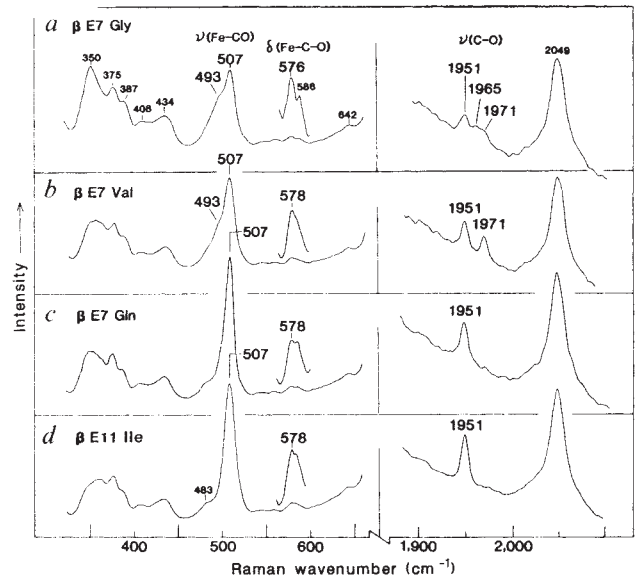


Fig. 3 Resonance Raman spectra of mutant Hbs in the carbon monoxy- form. *a*, $\alpha_2\beta$(His-E7→Gly)$_2$; *b*, $\alpha_2\beta$(His-E7→Val)$_2$; *c*, $\alpha_2\beta$(His-E7→Gln)$_2$; *d*, $\alpha_2\beta$(Val-E11→Ile)$_2$. The absence of a Raman line at 1355 cm^{-1} (data not shown) indicates that CO of the β *haem* is not photolysed by the excitation light and that both α and β subunits contribute to the spectra.

Methods. Resonance spectra of CO-Hb were recorded using a highly sensitive multi-channel Raman system[12]. A Spectra-Physics (Mountain View, California) Model 171-01 Krypton ion laser was employed for excitation at 406.7 nm. Laser power at the sample was maintained between 10 and 20 mW, and a spinning Raman cell was used to minimize photodissociation of CO and local heating. Wave numbers were calibrated using standard compound and spectra were recorded in 0.05 M Tris-HCl buffer, *p*H 8.0.

capable of donating a hydrogen bond to the oxygen molecule in a similar manner and the His-E7β to Gln substitution is therefore a structurally and functionally conservative change.

This paper is dedicated to Dr Max Perutz, who started this project 50 years ago. We are grateful to him for criticism and continuous support. We also thank Drs G. Fermi, J. Baldwin, D. Rhodes, D. Altschuh, R. Turner, J. C. Kendrew and J. Tame for helpful discussion, Dr A. Klug for interest and support, Dr P. Esnouf for factor X, Dr D. Bashford for globin sequence alignment, Colin Young for cell culture and Terry Smith for oligonucleotide synthesis. This work was supported by the MRC, NIH and the Veterans Administration. K.I. was supported by Ajinomoto.

Received 29 May; accepted 17 September 1987.

1. Phillips, S. E. V. *J. molec. Biol.* **142,** 531-554 (1980).
2. Shaanan, B. *J. molec. Biol.* **171,** 31-59 (1983).
3. Baldwin, J. *J. molec. Biol.* **136,** 103-128 (1980).
4. Brzozowski, A. *et al. Nature* **307,** 74-76 (1984).
5. Collman, J. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **71,** 1326-1329 (1974).
6. Peng, S. M. & Ibers, J. A. *J. Am. chem. Soc.* **98,** 8032-8036 (1976).
7. Phillips, S. E. V. & Schoenborn, B. P. *Nature* **292,** 81-82 (1981).
8. Perutz, M. F. *Nature* **228,** 726-739 (1970).
9. Nagai, K. & Thogersen, H. C. *Nature* **309,** 810-812 (1984).
10. Nagai, K., Perutz, M. F. & Poyart, C. *Proc. natn. Acad. Sci. U.S.A.* **82,** 7252-7255 (1985).
11. Luisi, B. & Nagai, K. *Nature* **320,** 555-556 (1986).
12. Tsubaki, M., Srivastava, R. B. & Yu, N.-Y. *Biochemistry* **21,** 1132-1140 (1982).
13. Fermi, G. *et al. J. molec. Biol.* **175,** 159-174 (1984).
14. Luisi, B. F. *thesis, Cambridge Univ.* (1986).
15. Petutz, M. F. & Mathews, F. S. *J. molec. Biol.* **21,** 199-202 (1966).
16. Kendrew, J. C. *Scient. Am.* **205,** 696-110 (1961).
17. Perutz, M. F. *J. crystal Growth* **2,** 54-56 (1968).
18. Jack, A. & Levitt, M. *Acta crystallogr., sect. A* **34,** 782-791 (1978).
19. Fermi, G. *et al. J. molec. Biol.* **155,** 495-505 (1982).
20. Monod, J., Wyman, J. & Changeux, J.-P. *J. molec. Biol.* **12,** 88-118 (1965).
21. Baldwin, J. M. *Prog. Biophys. molec. Biol.* **29,** 225-320 (1975).
22. Imai, K. *Allosteric Effects in Haemoglobins* (Cambridge University Press, 1982).
23. Makinen, M. W., Houtchens, R. A. & Caughey, W. S. *Proc. natn. Acad. Sci. U.S.A.* **76,** 6042-6046 (1979).
24. Hanson, J. C. & Schoenborn, B. P. *J. molec. Biol.* **153,** 117-146 (1981).
25. Stenzel, P. *et al. J. biol. Chem.* **254,** 2071-2076 (1979).